Fast computational treatment of molecular complexation in solution: Hybrid integral equation approaches

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigte
Dissertation

vorgelegt von

Dipl.-Ing. Friedemann Schmidt
aus München

Berichterstatter: Prof. Dr. J. Brickmann
Mitberichterstatter: Prof. Dr. H. J. Lindner
Tag der Einreichung: 24.09.2001
Tag der mündlichen Prüfung: 5.11.2001

Darmstadt 2001

D 17

The most exciting phrase in science,
the one, that heralds new discoveries,
is not 'Eureka!' (I found it!),
but 'That's funny.....'

Isaac Asimov

Für Jana
und meine Eltern

Diese Dissertation wurde am Institut für Physikalische Chemie an der Technischen Universität Darmstadt unter der Leitung von Prof. J. Brickmann in der Zeit von Januar 1998 bis September 2001 durchgeführt.
Die Arbeit wurde in englischer Sprache angefertigt. Eine Zusammenfassung der Methoden und Ergebnisse in deutscher Sprache findet sich im letzten Kapitel.

Durch das Wachstum des Informationsmarkts ist Kommunikation und Zusammenarbeit zu einem sehr wichtigen Bestandteil wissenschaftlichen Arbeitens geworden. Ich möchte mich bei all denjenigen bedanken, die auf die eine oder andere Weise zum Gelingen dieser Arbeit beigetragen haben:

Mein besonderer Dank gilt Prof. Dr. Jürgen Brickmann für alle unterstützenden und förderlichen Maßnahmen während meines Studiums, der Diplomarbeit und der Doktorandenzeit. Die Themenstellung der Arbeit kam meinen Interessen sehr entgegen, und so habe ich auch den mir gewährten Freiraum bei der Auswahl geeigneter Methodiken und der Verwirklichung des Arbeitsplans sehr genossen.

Auch Dr. Stefan Kast möchte ich an dieser Stelle hervorheben. Ich bedanke mich für alle Ideen, Anregungen und die zahlreichen Diskussionen rund um Computerchemie, Theorie und vielem anderen mehr, die bei der Verwirklichung der Arbeit von unschätzbarem Wert waren.

Ein ganz herzliches Dankeschön für hervorragende Zusammenarbeit geht an Dipl. Ing. Bernd Schilling und meine ehemaligen Kollegen Dr. Robert Jäger und Dr. Dirk Zahn. Die gemeinsame Zeit an der Hochschule war auch in sportlicher Hinsicht eine Herausforderung.

Bei Dr. Hans-Jürgen Bär, Dipl. Ing. Matthias Keil, Dr. Marco Müller, Dr. Thomas Exner und Dipl. Ing. Thorsten Borosch und den Mitgliedern des Arbeitskreises bedanke ich mich für das gute Arbeitsklima und vielerlei Hilfe beim täglichen gemeinsamen Kampf gegen den Computer.

Nicht zuletzt geht auch ein herzliches Dankeschön an Prof. Dr. Klaus-Peter Dinse, Ulla Henkes und die Mitglieder des Graduiertenkollegs für die gemeinsamen Veranstaltungen in Darmstadt und Hirschegg.

Danke!

Inhalt

1 Introduction

During recent years, the research activities in the fields of liquid phase *host-guest chemistry* have grown rapidly.[1] Much of this effort has been triggered by unique properties that are found within *supramolecular complexes*. Some of these systems exhibit specific patterns of complementary functionality to enable recognition processes such as protein-inhibitor binding or selective complexation of atomic or molecular guests.[2] Sometimes the binding constant of a certain complex exceeds the values of very similar ones by several orders of magnitude. These hosts can be termed specific for a molecular guest. Specific interactions play an important role in nature, including information storage and transfer, catalysis, self-assembly and molecular recognition.

A broad spectrum of binding modes has been found, which depends on the nature of the complex partners. Host-guest interaction is sometimes dominated by electrostatics or hydrophobic interaction, hydrogen bonding or shape complementary. Beyond it, steric effects are of major importance in recognition phenomena. Some of these qualities are useful for the development of purification and separation methods as well as selective catalyst research in the field of technical and synthetic chemistry.[3]

By this means, crown ethers such as 18-crown-6 have been used for a long time[4] as selective complexing agent for metal ions and amines in organic synthesis. However, the binding modes of atomic and molecular guests are substantially different. While metal ions show a tendency to form inclusion complexes,[5] the binding of organic amines is often controlled by hydrogen bonding.[1]

Particularly the production of pharmaceuticals critically depends on the availability of pure reactants and products. For practical purposes, it is often necessary to produce enantiopure compounds, which can generally be achieved by interaction of the mixture with a chiral selector[6] in the sense of diastereomeric host-guest complexes. Chiral crown ethers[7,8] and cyclodextrines[9,10] have proved very successful in the enantioseparation of organic amines and other substrates.

Among other quantities, thermodynamics and structural parameters have been extensively measured to understand interactions within host-guest complexes in solution. The thermodynamic stability of a given complex is classified by its free

energy of complexation. Accordingly, the relative stability of a set of complexes can be identified with the difference of their free energies of complexation. Both experimental and computational chemistry allow the determination of thermodynamic parameters of complex formation.

The fast, but accurate calculation of free energies of complexation is a major challenge in the computational treatment of molecular recognition. Consequently, a variety of methods have been developed over the last decades in order to calculate the thermodynamics of complexation in the liquid phase. These methods mainly differ with respect to the treatment of the solvent phase. Most of these methods fall either under the category of *empirical approaches* or the category *of simulation approaches*.

Empirical approaches often use implicit models of the solvent phase, in which the bulk solvent phase is treated as a structureless continuous medium. Implicit solvent models often yield accurate electrostatic free energies[11,12] with a reasonable computational effort. However, the neglect of the solvent structure is not appropriate whenever specific solute-solvent interactions such as hydrogen bonding occur in solution.

Molecular dynamics (MD) or *Monte-Carlo* (MC) simulations[13] with explicit solvent molecules and periodic boundary conditions result a detailed picture of the solvent structure and dynamics. But, a general shortcoming of any atomistic treatment of the solvent phase is the large effort spent for sufficient sampling of the solvent phase space.

Hybrids of empirical and simulation methods allow to combine some of their individual advantages. Hybrid algorithms offer an favorable base to the reduction of the CPU time demand of molecular simulations while retaining universality and the prediction power of classical simulation methods. Therein, solute-solute interactions are treated by simulation methods, while the solvent is modeled implicitly. Within the statistical-mechanical *integral equation theory*,[14] the molecular nature of the solvent is simplified by means of precomputed solvent-solvent radial distribution functions. Thermodynamic properties of solution can be computed quickly using the *reference interaction site model* (RISM).[15,16] On the one hand, computation times can be reduced by several orders of magnitude as compared to atomistic *molecular dynamics* (MD) or *Monte-Carlo* (MC) simulations. On the other hand, important information about the spatial structure and the dynamics of the system is lost, and

can not be reconstructed from the set of distribution functions. However, RISM integral equation theory was successfully employed in conjunction with well-established computational techniques to study intramolecular *potential energy surfaces* of polymers[17] in solution.

It is the central intention of this thesis to combine the RISM integral equation model with sampling and optimization strategies for solute-solute interactions in order to explore *free energy surfaces* (FES) of molecular association. This work is organized as follows:

In section 2, a hybrid integral equation/Monte Carlo RISM/MC algorithm is presented, which allows for fast calculation of crown ether - metal ion thermodynamics in solution. Calculation times commonly spent in sampling the solvent configurational space are reduced drastically by the implementation of an interaction site model of the solvent. This statistical model is easily transferable to a number of solvents, which can be described by classical force fields.[18,19,20] Solute-solute statistics are sampled by use of *statistical perturbation theory*.[21] Therein, the well-established *hypernetted chain closure* (HNC)[14] is used to derive solvent contributions to the free energy of the system. This is accomplished by solving the RISM integral equations once for a constrained supramolecular complex and using a variational free energy expression for conformational changes.[22]

The construction of free energy surfaces of rigid molecules with this variational free energy expression is straightforward. For the complex formation of alkali metal ions Na^+, K^+, Rb^+ and Cs^+ with 18-crown-6 in methanol and acetonitrile, orientationally averaged *potentials of mean force* (PMF) are presented. A radial integration procedure is employed to extract standard free energies of complexation $\Delta G°$ from the PMF.

In chapter 3, an extension of the RISM/MC approach is presented. The improved algorithm allows for the calculation of the complexation thermodynamics of hydrogen bonding molecules. It is shown that the integral equation approach is capable to provide essential features of the solvation structure around the host 18-crown-6. The FES of $H_3CNH_3^+$/18-crown-6 in water and methanol are studied in detail. The anchoring mechanism of urea/18-crown-6 and some guanidinium derivatives in methanol is discussed by means of specific hydrogen bond formation. In complexes with the smaller 15-crown-5, the same guests form two hydrogen bonds that can be characterized in a similar way.

Orientationally averaged potentials of mean force are presented for the hydrogen-bonded crown ether complexes. Free energies of complexation are presented for the various complex partners. The results are corrected for standard state conditions and compared with experimental data.

In chapter 4, the hybrid integral equation genetic algorithm method for molecular docking is introduced and applied to chiral selection. A genetic algorithm in combination with integral equation theory (RISM/LGA) is employed to explore the local minima of the FES with respect to locating the global minimum. Therein, constrained local optimization is used to enhance the performance of the genetic algorithm. The relative stability of enantiomeric host-guest complexes termed chiral selection, is expressed as the difference of the global minima of the FES, $\Delta(\Delta G_{R\text{-}S})$, between enantiomeric complexes and is compared with experimental data.

1.1 References

(1) Lehn, J.-M. Supramolecular Chemistry: Concepts and Perspectives, VCH, Weinheim **1995**.

(2) Westman, G. Selectivity in Host-Guest Chemistry, Thesis, Department of Organic Chemistry, Chalmers **1995**.

(3) Lide, D.R.; Ed.; CRC Handbook of chemistry and physics, CRC Press **1994**.

(4) Petersen, C.J. *J. Am. Chem. Soc.* **1967**, *89*, 7017.

(5) Vögtle, F. Supramolekulare Chemie, Teubner Verlag, Wiesbaden **1992**.

(6) Rudolph, J. Selektive Erkennung von Molekülen und stereotypen funktionellen Gruppen, Thesis, Bochum **1995**.

(7) Stoddard, J.F.; Eliel, E.L.; Wilen, S.H. Topics in stereochemistry, Wiley Interscience New York, **1988**.

(8) Zhang, X.X.; Bradshaw, J.S.; Izatt, R.M. *Chem. Rev.* **1997**, *97*, 3313.

(9) Takahashi, K.; Hattori, K. *J. Inclus. Phenom. Mol. Recog. Chem.* **1994**, *17*, 1.

(10) Lipkowitz, K.B., *Chem. Rev.* **1998**, *98*, 1741.

(11) Nicholls, A.; Honig, B. *J. Comp. Chem.* **1991**, *12*, 435.

(12) Schäfer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100*, 1578.

(13) Allen,M.P.; Tildesley, D.J. Computer Simulation of Liquids, Claredon Press, Oxford **1987**.

(14) Hansen, J.P.; McDonald, I.R. *Theory of simple liquids*, Academic Press, London **1980**.

(15) Chandler, D.; Andersen, H.C. *J. Chem. Phys.* **1972**, *57*, 1930.

(16) Hirata, F.; Rossky, P.J. *Chem. Phys. Lett.* **1981**, *83*, 329.

(17) Khalatur, P. G.; Khokhlov, A. R. *Mol. Phys.* **1998**, *93*, 555.

(18) Jorgensen, W.L. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(19) Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M. Jr.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(20) MacKerell, A.D.; Bashford, D.; Bellott, M.; Dunbrack, R.L.; Evaseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; McCarthy, D.; Kucnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D.T.; Prohom, B.; Reiher, W.E.; Roux, B.; Schlenkrich, M.; Smith, J.C.; Stote, R.; Straub, J.W.; Tanabe, M.; WiorkiewiczKuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(21) Prue, J.E. *J. Chem. Educ.* **1969**, *46*, 12.

(22) Singer, S.J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

2 The Hybrid Integral Equation/Monte Carlo Approach to Complexation Thermodynamics

2.1 Summary

A new method for the rapid calculation of complexation free energy surfaces (FES) in solution is presented. The Monte-Carlo (MC) simulation technique for solute-solute statistics is combined with the reference interaction site model (RISM) integral equation theory for solute-solvent and pure solvent statistics. Statistical-mechanical perturbation theory is used to determine complexation potentials of mean force and, from these, standard free energies of binding. The hybrid RISM/MC method is applied to calculate the selectivity of rigid 18-crown-6 for alkali metal ions in liquid methanol and acetonitrile. A full atom model of the host in its D_{3h} conformation is used for the calculations. For the standard free energies of binding good agreement with the experimental data is found. Additional information about the structure of the FES is extracted by scanning its topography. Two-dimensional slices through the FES are presented to illustrate the solvation structure of the complexes. By connecting local minima and saddle points, minimum energy pathways are revealed, leading to the globally optimal complex structure.

2.2 Introduction

The accurate calculation of free energies of complexation is one of the most challenging problems in the computational treatment of molecular recognition. Much effort has been spent during the past years to develop and improve microscopically detailed simulation techniques for the accurate calculation of binding affinities.[1] A general shortcoming of an atomistic treatment of the solvent phase is the large computational power needed for sufficient sampling of the solvent molecules which motivated the development of approximate solvent models. Thermodynamic properties of solution systems can be computed quickly with sufficient accuracy by using approximate solvent models. Implicit models of the solvent simplify the molecular nature of the liquid by continuum properties like the dielectric constant

while the more detailed statistical solvent models replace the solvent by distribution functions that can be computed ahead of treating the solute degrees of freedom. A speedup of several orders of magnitude in computation time as compared to atomistic molecular dynamics or Monte-Carlo (MC) simulations is therefore possible. Roux and Simonson[2] give a review of this topic.

The reference interaction site model (RISM) integral equations yield a statistical solvent model comprising a detailed description of the solvent structure depending on the conformation of solute molecules. Although the distribution functions are approximate, the structural results were shown to agree reasonably well with simulation data for pure liquids[3,4,5] and for solutions.[6,7,8,9,10,11] Since the free energy of solvation can be determined quickly from the solution of the integral equations the solvent-induced potential of mean force (PMF) for a given solute geometry is readily accessible. Further reduction to simpler intra-solute coordinates can be achieved by averaging over the resulting PMFs, usually by sampling intramolecular degrees of freedom from molecular simulations. This approach has been applied to polymer systems (PRISM/MC)[12,13] and to conformational MC sampling of single biomolecules.[14,15] In certain circumstances intra-solute correlation functions and, from these, the PMFs can be determined from a suitable integral equation alone: The PMFs for simple ions in solution have been calculated in the past utilizing both the one-dimensional[16] (1D) as well as the three-dimensional[17,18] (3D) RISM approach. Orientationally dependent PMFs of molecular ions in aqueous solution have been determined from the 3D-RISM approach.[19] This formally provides a route to free energies of binding via averaging over all orientations, but the computational effort is however large.

Although the shortcomings of the 1D-RISM treatment as compared to the 3D case are well known[20] the former have a tremendous computational advantage when combined with a MC treatment of certain degrees of freedom, as will be shown later. The value of this procedure can be assessed by comparison with available experimental data. To this end, we present in this paper a method that combines MC sampling of the solute-solute configurational space for bimolecular association with the 1D-RISM equations within the hypernetted chain (HNC) approximation, applied to the complexation of alkali metal ions and 18-crown-6 in methanol and acetonitrile. Several aspects of complexation will be discussed: 2D scans and minimum energy paths (MEP) of the free energy surface (FES) or multidimensional PMF reveal its

topography and relevant mechanisms of binding in solution, free energy perturbation calculations yield effective 1D PMFs from which the free energy of complexation is determined. The article is outlined as follows: In the next section the hybrid MC/RISM approach is described. Section 3 provides computational details of the methods used. Our results are reported in section 4 and summarized in the final section.

2.3 Theory

2.3.1 RISM Equations and Free Energy Determination.

The 1D-RISM equations, initially developed by Chandler and Andersen,[21] are nonlinear integral equations for the correlation functions of molecular liquids composed of sites α and γ taking the form

$$\mathbf{h}^{vv} = \boldsymbol{\omega}^{v} * \mathbf{c}^{vv} * \left(\boldsymbol{\omega}^{v} + \rho\mathbf{h}^{vv}\right), \tag{2.1}$$

$$\mathbf{h}^{Xv} = \boldsymbol{\omega}^{X} * \mathbf{c}^{Xv} * \left(\boldsymbol{\omega}^{v} + \rho\mathbf{h}^{vv}\right), \tag{2.2}$$

$$\mathbf{h}^{YX} = \boldsymbol{\omega}^{Y} * \mathbf{c}^{YX} * \boldsymbol{\omega}^{X} + \boldsymbol{\omega}^{Y} * \mathbf{c}^{Yv} * \rho\mathbf{h}^{Xv}, \tag{2.3}$$

where $\mathbf{h} = (h_{\alpha\gamma}(r))$ is the matrix of the total correlation functions (the radial distribution function is $g = h+1$), $\mathbf{c} = (c_{\alpha\gamma}(r))$ are the direct correlation functions, $\boldsymbol{\omega} = (\omega_{\alpha\gamma}(r))$ are the intramolecular correlation functions (for rigid molecules these are normalized Dirac delta functions constraining the site distance), ρ is the site density, r is the distance and the asterisk denotes convolution. Here, eq 2.1 is valid for pure solvents (superscript v), eq 2.2 is the solute-solvent relation in the limit of infinite solute dilution (superscript X or Y), eq 2.3 yields the correlation functions between two solutes X and Y. The integral equations must be supplied with a suitable closure relation. Various closures have been proposed,[22,23,24] among them particularly for polar systems the HNC approximation[25]

$$h_{\alpha\gamma} = \exp\left(-\beta u_{\alpha\gamma} + h_{\alpha\gamma} - c_{\alpha\gamma}\right) - 1 \tag{2.4}$$

provides the best approximation to the pair distribution functions. Here, $u_{\alpha\gamma}$ is the pair potential and $\beta = 1/kT$ where k is the Boltzmann constant and T is the absolute temperature. The coupled eqs 2.1-2.3 and 2.4 are solved iteratively. The pair

potentials used in this work are the sum of a short-ranged Lennard-Jones term and a long-ranged Coulomb term

$$u_{\alpha\gamma} = u_{\alpha\gamma}^{LJ} + u_{\alpha\gamma}^{C} = 4\pi\varepsilon_{\alpha\gamma}\left[\left(\frac{\sigma_{\alpha\gamma}}{r_{\alpha\gamma}}\right)^{12} - \left(\frac{\sigma_{\alpha\gamma}}{r_{\alpha\gamma}}\right)^{6}\right] + \frac{q_\alpha q_\gamma}{r_{\alpha\gamma}}, \tag{2.5}$$

where q_i is the partial atomic charge on site *i*. For the Lennard-Jones parameters $\varepsilon_{\alpha\gamma}$ and $\sigma_{\alpha\gamma}$ the standard Lorentz-Berthelot mixing rules are used. To insure dielectric consistency at least on the level of a phenomenological dielectric constant, for the calculation of the pure solvent correlation functions the solvent site charges are scaled so that the correct asymptotic behavior of the direct correlation functions is obtained.[16,26]

With a given set of HNC solute-solvent correlation functions, the excess chemical potential or transfer free energy from gas phase to solution is given[27] in closed form as

$$\Delta\mu = -\frac{\rho}{\beta}\sum_{\alpha,\gamma}\int d\mathbf{r}\left(\frac{1}{2}h_{\alpha\gamma}^2 - c_{\alpha\gamma} + \frac{1}{2}h_{\alpha\gamma}c_{\alpha\gamma}\right). \tag{2.6}$$

Information about the solute-solute correlation can in principle be extracted from eq 2.3. Unfortunately, iterations of eq 2.3 diverge when applied to pairs of multi-site molecules. Consequently, eq 2.3 has so far been solved for pairs of monoatomic solutes[16,17,18] only. The 3D generalization[19] of the solute-solute equation can be solved for fixed molecular orientations only. To overcome this problems with eq 2.3, a superposition approximation[28] was proposed where the PMF is assumed to be pair-decomposable, i.e. solute molecules are split into atomic sites, interatomic correlation functions are obtained from eq 2.3, and the full PMF is constructed according to a linear superposition of the site-site PMFs.

Although this approach has been successfully used in the construction of an intramolecular solvation FES,[29,30] we do not adopt the superposition approximation in this work. Instead, the orientationally dependent PMF *W* between solute molecules of arbitrary shape can be readily calculated via

$$W(\mathbf{\Gamma}) = -\Delta\mu_{\mathrm{X}} - \Delta\mu_{\mathrm{Y}} + \Delta V_{\mathrm{XY}}(\mathbf{\Gamma}) + \Delta\mu_{\mathrm{XY}}(\mathbf{\Gamma}). \tag{2.7}$$

according to the thermodynamic cycle in Figure 1.

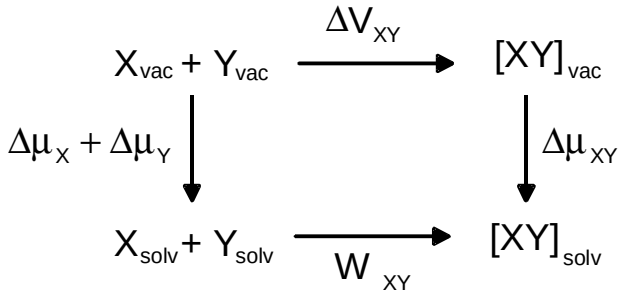

Figure 1. Thermodynamic cycle of the supramolecular approach for the indirect determination of free energies of complexation in solution.

This corresponds to the construction of a supermolecule XY constrained by a given set of six intermolecular coordinates $\mathbf{\Gamma} = \{r, \theta, \varphi, \Theta, \Phi, \Psi\}$, thus defining a FES where the solvation free energy and solute-solute interaction is taken into account. On the right hand side of eq 2.7, only the vacuum interaction ΔV_{XY} and the supermolecule chemical potential $\Delta\mu_{XY}$ depend on the intermolecular coordinates. ΔV_{XY} is calculated by eq 2.5, while eq 2.6 is used for the calculation of free energies of solvation. If we consider one of the complex partners to be a monoatomic solute Y eq 2.7 simplifies to

$$W(r,\theta,\varphi) = -\Delta\mu_X - \Delta\mu_Y + \Delta V_{XY}(r,\theta,\varphi) + \Delta\mu_{XY}(r,\theta,\varphi) \tag{2.8}$$

where r, θ, φ are spherical polar coordinates specifying the single site solute position for instance with respect to the host center of mass. The FES is then spanned in the space of the three coordinates.

Eq 2.8 in principle requires to solve the RISM equations for every intermolecular configuration which is a time-consuming process due to the iterative solution technique. However, Ref.[31] provides a simple expression which relates the change in the free energy of solvation upon a conformational variation $\{r, \theta, \varphi\} \rightarrow \{r', \theta', \varphi'\}$ to a simple difference of the intramolecular correlation matrices by

$$\Delta\mu(r',\theta',\varphi') - \Delta\mu(r,\theta,\varphi) = -\frac{\rho}{2\beta}\sum_{\alpha,\alpha'}\int d\mathbf{r}\left[\mathbf{c}^{Xv} * \left(\boldsymbol{\omega}^{v} - \rho\mathbf{h}^{vv}\right) * \mathbf{c}^{Yv}\left(\boldsymbol{\omega}^{XY'} - \boldsymbol{\omega}^{XY}\right)\right]_{\alpha,\alpha'} . \tag{2.9}$$

This equation is derived within the HNC approximation by writing down the total variation of $\Delta\mu$ with respect to **c**, **h**, **h-c**, and **ω**, and applying the variational principle.[27] Here, the summation has to performed over all sites $\{\alpha;\alpha'\}$ of the

molecules X and Y. Explicit RISM-HNC solutions are therefore needed only for one single configuration, the remaining solvent-induced part of the FES is constructed by simple matrix multiplication (of the Fourier transforms of the quantities under the integral) and integration, without the need for an iterative process. The use of eq 2.9 tremendously reduces computation times, particularly if used within an MC procedure.

2.3.2 RISM/MC

From eqs 2.7, 2.8, and 2.9 the FES of the interaction of solvated rigid molecules can be determined. However, even if all intramolecular degrees of freedom are frozen, a computational scan of the FES is still expensive. On a rectangular grid of 40 Å with a spacing of 0.1 Å this requires approximately 10^8 evaluations. On the other hand, the calculation of 2D slices of the FES illustrating topographical features can be done quickly. The computation of a one-dimensional PMF $w(r)$ from the multi-dimensional W requires a special sampling procedure. To this end, we have implemented a MC procedure based on statistical-mechanical perturbation theory.[32,33,34,35] The PMF difference that results from a structural change is given by

$$w(r') - w(r) = -kT \ln\left\langle \exp\{\beta(W(r,\theta,\varphi) - W(r',\theta,\varphi))\}\right\rangle_r . \qquad (2.10)$$

For our purposes, the average represents the sampling of angular coordinates with respect to a small change in the reaction coordinate r. Eq 2.10 gives only free energy differences between states r and r'. The PMF offset is calculated from eq 2.8 at such a large distance that the solvent-mediated interaction approaches zero.

A flow chart for the hybrid RISM/MC algorithm reads as follows:

I. Eqs 2.1, 2.2, 2.4, and 2.6 are solved to yield free energies of solvation of the complex partners.

II. A supramolecular structure is constructed by placing X and Y at a large distance r. Eqs 2.2, 2.4, and 2.6 are solved to obtain a set of reference distribution functions **h**, **c**(r_o). The offset of the PMF is calculated from eq 2.8.

III. A set of host-guest configurations is generated randomly by variation of the angular coordinates θ and φ. The PMF difference $W(r',\theta,\varphi)$-$W(r,\theta,\varphi)$ for a small step of the guest towards the host is calculated via eqs 2.8 and 2.9.

IV. Configurations are accepted with the probability P meeting the Metropolis criterion

$$P = \begin{cases} 1 & \Leftrightarrow W(r,\theta,\varphi') \geq W(r,\theta,\varphi) \\ \exp\{\beta(W(r',\theta,\varphi) - W(r,\theta,\varphi))\} & \Leftrightarrow W(r',\theta,\varphi) < W(r,\theta,\varphi) \end{cases} \tag{2.11}$$

V. The free energy difference $w(r')$-$w(r)$ for a step of the guest towards the host is sampled according to eq 2.10.

Steps III-V are repeated until the reaction coordinate approaches zero.

The resulting PMF contains the information about the free energy surface which is essential to calculate free energies of binding of the complex partners. We use the integration formula presented by Pranata and Jorgensen[34] which relates the one-dimensional PMF to an association constant of a host-guest complex:

$$\Delta G_{\mathrm{XY}} = -kT \ln\left\{ \frac{4\pi}{V_0} \int_0^{r_c} dr\, r^2 \exp(-\beta w(r)) \right\}. \tag{2.12}$$

Here, a change of volume-pressure work is assumed to be negligible. All bound states are considered to contribute to the overall binding free energy, where the complex partners move within a contact distance in the first energy minimum. The integration therefore runs from zero to a certain cutoff distance r_c where the contact minimum ends. Contributions of complex partners moving along outer solvent-induced energy wells are neglected. This equation however does not account for the thermodynamic standard state defined by a concentration of 1 mol l^{-1} as used in experimental work, nor does it consider changes of rotational symmetry. All these corrections are readily formulated.[36,37,38,39] In case of ion-crown ether interaction there is no change in the rotational symmetry of the complex with respect to the unbound state.[40] Although eq 2.12 has been used in the past without further correction[35,41] the correct form of the standard free energy of complexation in our case reads

$$\Delta G^{\mathrm{o}}_{\mathrm{XY}} = \Delta G_{\mathrm{XY}} - RT \ln(C^{\mathrm{o}} V^{\mathrm{o}}) = \Delta G_{\mathrm{XY}} - 4.38 \text{ kcal mol}^{-1} \tag{2.13}$$

where R is the gas constant and V^{o} equals to 1 $Å^{-3}$.

2.3.3 Computational Details

We used the *optimized potential for liquid simulation* (OPLS)[42,43] that has been extensively used for MC simulations of liquid systems. The nonbonded interaction parameters for ions were taken from ref. [44]. OPLS provides three-site united-atom models for both methanol and acetonitrile. Geometric parameters as well as partial charges and Lennard-Jones parameters for acetonitrile and methanol were taken directly from OPLS with the exception of the hydoxyl H of methanol. To avoid Coulomb singularities which arise inevitably in the convolution of distribution functions when applied to nonrepulsive interaction sites, we set $\sigma_{HO} = 0.4$ Å and $\varepsilon_{HO} = 0.046$ kcal mol^{-1}. The structure and nonbonded parameters of the crown ether in its D_{3h} symmetric conformation were also taken from the OPLS force field. The crown ether was kept rigid during the calculations. A complete overview of the nonbonded parameters used in this study is given in Table 1.

TABLE 1: Force field parameters: atomic partial charges *q*; Lennard-Jones parameters σ and ε

Site	*q* / e	σ / Å^{-1}	ε / kcal mole^{-1}
Methanol O	-0.7000	3.070	0.1700
Methanol CH_3	0.2650	3.775	0.1180
Methanol H	0.4350	0.400	0.0460
Acetonitrile C	0.2800	3.650	0.1500
Acetonitrile CH_3	0.1500	3.775	0.2070
Acetonitrile N	-0.4300	3.200	0.1700
18-crown-6 C	0.1400	3.500	0.0660
18-crown-6 O	-0.4000	2.900	0.1400
18-crown-6 H	0.0300	2.500	0.0300
Na^+	1.0000	2.85	0.0314
K^+	1.0000	3.56	0.1304
Rb^+	1.0000	3.92	0.1530
Cs^+	1.0000	4.29	0.3594

The RISM-HNC equations were solved on a logarithmically spaced grid of 512 points, ranging from a minimum of $5.98 \cdot 10^{-3}$ Å to a maximum of 164.02 Å. For the calculation of fast fourier transforms we used the Talman procedure.[45,46] Number densities of methanol and acetonitrile at 298.15 K were taken as 0.01405 $Å^{-1}$ and 0.0116 $Å^{-1}$ as suggested for to liquid state simulations by Jorgensen,[47] the dielectric constants were set to 32.66 for methanol and 35.94 for acetonitrile. The solvent-solvent and solute-solvent integral equations were iteratively converged by standard methods until the convergence criterion $\max(\Delta c_{\alpha\gamma}) < 10^{-5}$ was met.

2.4 Results

2.4.1 Structure of the FES

From the RISM solutions for alkali ions Na^+, K^+, Rb^+, Cs^+ and 18-crown-6 in methanol and acetonitrile the FES was constructed according to eqs 2.8 and 2.9. Characteristics of the FES can be visualized by two-dimensional slicing planes defined in a crown ether coordinate system depicted in Figure 2.

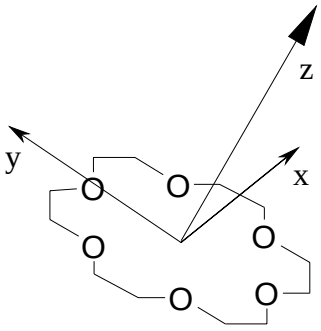

Figure 2. Coordinate system used for 18-crown-6. The ring plane of the crown is defined by the *x* and *y* axes.

A slicing plane through the FES of the system 18-crown-6-ion in methanol is presented in Figure 3. The slicing plane is equal to the *xz* plane given in Figure 2 that is perpendicular to the ring plane of the crown.

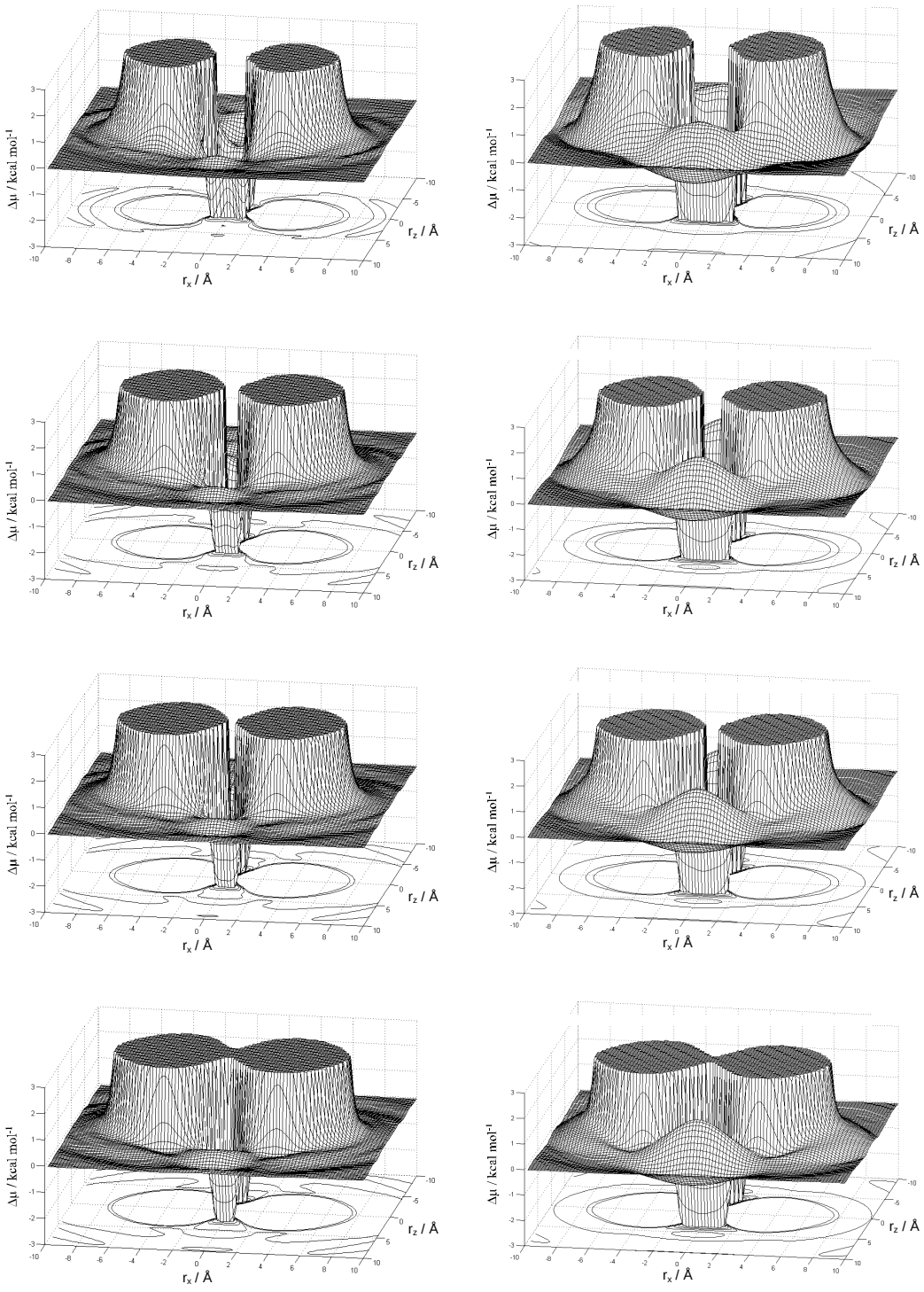

Figure 3. Slices through the FES of ion-18-crown-6 in methanol; left: methanol, right: acetonitrile; from top to bottom: Na^+, K^+, Rb^+, Cs^+. The *z* axis and *x* axis of the crown ring as given in Figure 2 point towards the origin of the *x* axis and *z* axis, respectively.

In case of K^+ and Na^+, a single contact minimum is found at the center of the crown, shielded by the crowns methylene groups. For Rb^+ and Cs^+, a strong barrier arises at the center of mass of the crown, prohibiting the passage of the ion through it. Approaching the center of mass of the crown, the ion must cross a region where the solvent shells of the host and the ion interfere. Oscillations are observed that depend strongly on the solute charges and the polarity of the solvent. In methanol, these oscillations are found to be of the order of 0.5 *kT*, while the local barriers in acetonitrile range from 1 to 2 *kT*.

The largest barrier for the ions is found to be raised by the ether oxygens along the *z* axis of the crown near the center of mass. Local interference of the ions with solvent molecules interacting with the crown ether oxygens generates a barrier of approximately 1-2 *kT* which prevents a collinear attack of the ion in solution. This barrier is most probably due to an acetonitrile located along the *z* axis of the crown. The relatively low central barrier in methanol refers to laterally orientated solvent molecules.

2.4.2 Potentials of Mean Force

Radial PMFs were derived by the hybrid RISM/MC free energy perturbation method. The reaction coordinate was defined as the center of mass-separation between the crown and the metal ion. The initial distance of the complex partners was chosen as 20 Å, insuring independence of the PMF of the initial configuration . 200 perturbation windows were sampled along the reaction coordinate, each with an offset of 0.1 Å. For each window, 10000 solute configurations were sampled yielding the average free energy difference for a 0.1 Å step of the ion towards the host. The results are depicted in Figure 4.

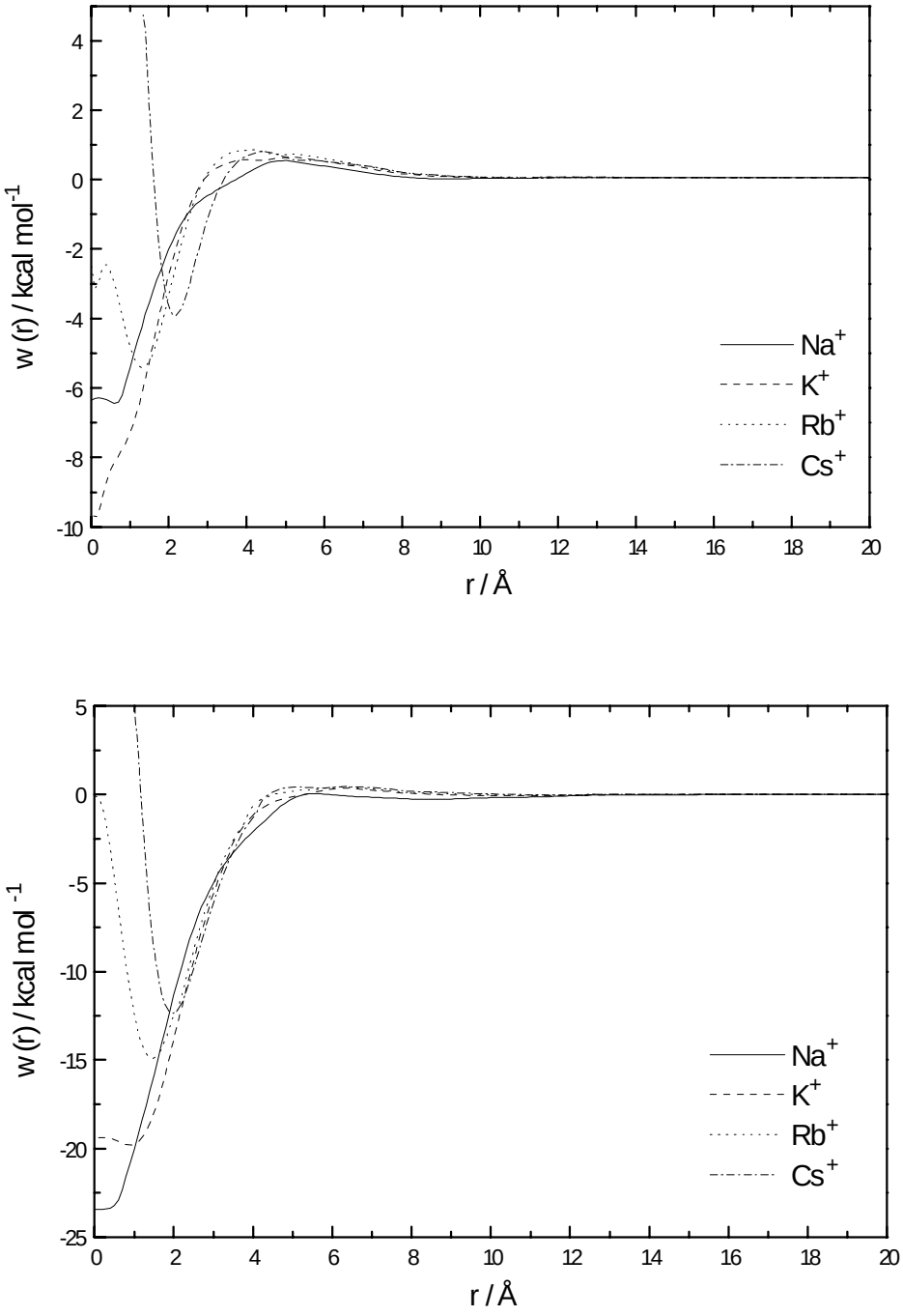

Figure 4. Orientationally averaged PMF for the host-guest interaction; top: in methanol, bottom: in acetonitrile. The reaction coordinate is given by the distance between the crown center of mass and the ion.

It is found that, depending on the size of the cation, the minima of the PMF deviate significantly from the center of mass of the crown. They are found at 0.2 Å, 1.0 Å, 1.9 Å, and 2.4 Å for Na^+, K^+, Rb^+ and Cs^+, respectively. In case of Na^+, which is smaller than the average cavity diameter of 5.6 Å, the ion prefers a slightly asymmetric position in the rigid crown that allows the ion to approach the three ether oxygens. In case of the larger ions, we can attribute this behavior to the steric demand of the ion. A similar situation for the K^+ ion in water has been reported earlier.[48] The depth of the minima in methanol varies as $K^+ > Na^+ > Rb^+ > Cs^+$ while in case of acetonitrile, we find $Na^+ > K^+ > Rb^+ > Cs^+$.

Eq 12 was used to compute free energies of binding from the one-dimensional PMFs. A cutoff distance of 5 Å was used throughout; test calculations showed that for these systems all integrations converge after a few Å and do not change much upon variation of the cutoff distance. The experimental and calculated free energies of binding are listed in Tables 2a and 2b. In general they show the same trends as the minima of the PMF. Good agreement of standard free energies with the experimental data[49,50] is found for the complexes in methanol, while those for acetonitrile tend to be overestimated.

TABLE 2a: Free energies of complexation in methanol

Ion	**Na^+**	**K^+**	**Rb^+**	**Cs^+**
Contact minimum / Å	0.2	1.0	1.9	2.4
Dissociative limit / Å	5.2	5.2	5.2	5.2
Exp. ΔG^o / kcal mol^{-1}	-6.4 ± 1.0	-8.3 ± 0.2	-7.4 ± 0.4	-6.4 ± 0.2
min($w(r)$) / kcal mol^{-1}	-6.29	-9.78	-5.42	-3.95
calc. ΔG / kcal mol^{-1}	-3.88	-4.28	-3.97	-3.87
calc. ΔG^o / kcal mol^{-1}	-8.26	-8.66	-8.35	-8.25

Apparently the free energies without standard state correction match the experimental data better in case of acetonitrile. This is however a mere conincidence. It is interesting to note though that the selectivity trend for a specific solvent is well represented and there is an almost constant offset of experimental vs. calculated

standard free energies of complexation in case of acetonitrile. This is probably related to the rigid geometry of the crown ether that it is taken as identical for both solvents. Furthermore the solvent model and the neglect of counterions could pose problems. The predictive relative capacity of the RISM/MC model for a given series in a specific solvent is however good.

TABLE 2b: Free energies of complexation in acetonitrile

Ion	**Na^+**	**K^+**	**Rb^+**	**Cs^+**
Contact minimum / Å	0.1	0.9	1.7	2.0
Dissociative limit / Å	5.0	5.0	4.5	4.5
Exp. ΔG^o / kcal mol^{-1}	-5.9 ± 0.8	-8.0 ± 0.5	-6.8 ± 0.4	-6.6 ± 0.2
min($w(r)$) / kcal mol^{-1}	-23.2	-19.9	-14.8	-12.6
calc. ΔG / kcal mol^{-1}	-8.59	-8.13	-6.80	-6.27
calc. ΔG^o / kcal mol^{-1}	-12.96	-12.50	-11.17	-10.64

2.4.3 Minimum Energy Path.

Insight into reaction pathways and complexation mechanisms can be gained from local minimization of a number of configurations subject to a distance constraint. This has been done with an angular downhill simplex minimizer that does not require gradient information.[51]

To insure smoothness of the path, the choice of the constrained distances is critical. We used a 0.1 Å spacing which is (see also Figure 3) considerably smaller than the average width of the solvent-induced barriers that are more than an order of magnitude larger than the spacing. Interconnections between these minimized points define the MEPs. This strategy allows a graphical representation of valleys leading to the local minima of the FES and is exemplarily depicted for acetonitrile in Figure 5. The picture for methanol is qualitatively very close to that of acetonitrile, probably due to the similar dielectric solvent properties, and not shown here. Visualization was performed with the MOLCAD package.[52,53]

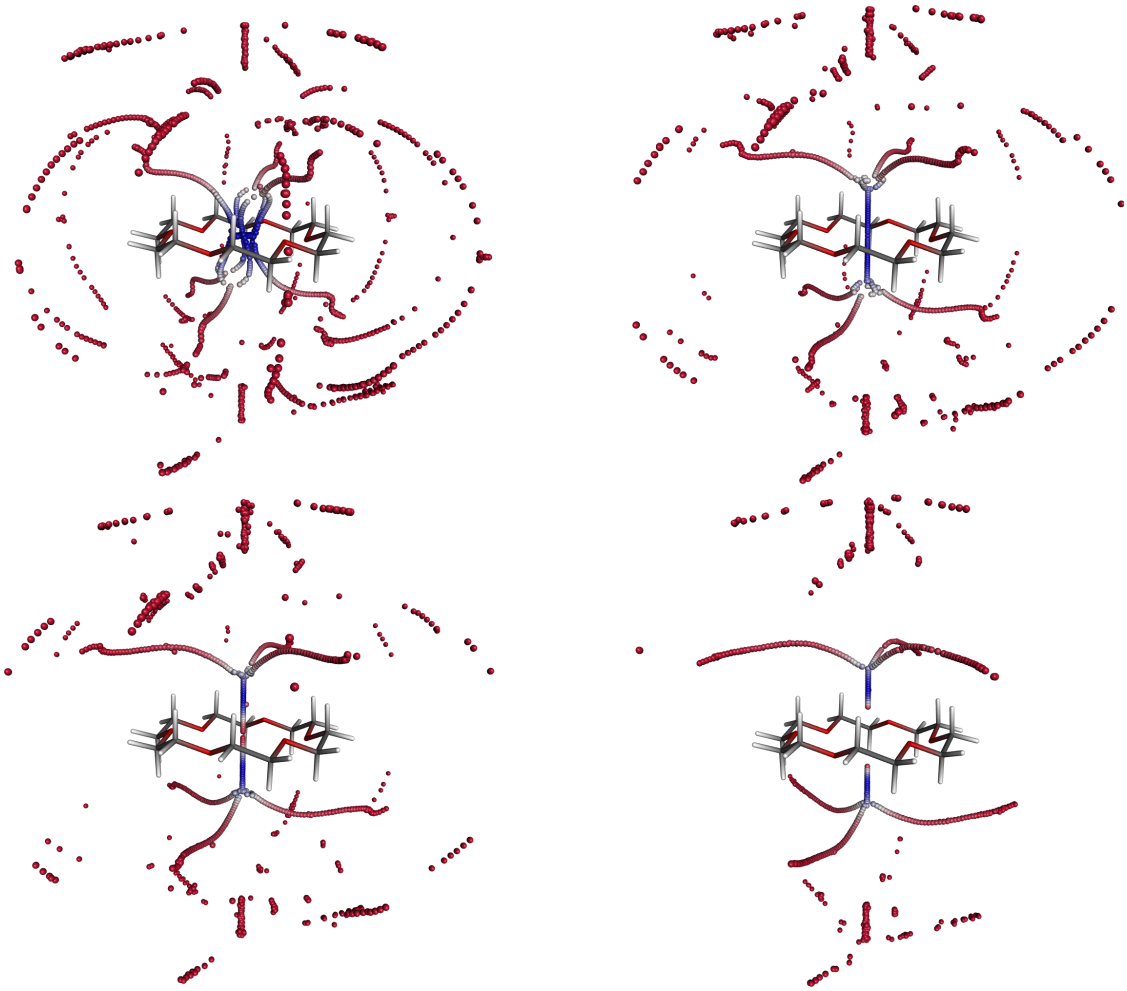

Figure 5. Color-coded minimum energy paths of Na^+, K^+, Rb^+ and Cs^+ approaching 18-crown-6 in acetonitrile. The scale ranges from blue (favorable) over white to red (unfavorable).

For the system K^+-18-crown-6 in aqueous solution, it has been stated that the ion's path approaching the host deviates strongly from the three-fold crown symmetry axis.[48] For electrostatic reasons, this behavior might also be expected from the vacuum-PES of potassium-18-crown-6.[48,54] However, in both methanol and acetonitrile a collinear attack of the potassium ion is preferred up to a distance of approximately 5 Å only. Climbing a small barrier, the ion is strongly attracted by a lateral oxygen. Therefore, it first moves to a lateral position near one ether oxygen, then proceeds directly to the three-fold symmetry axis.

In summary, two main effects can be attributed to the solvent influence: (1) Due to necessary desolvation of both host and guest molecules the overall attraction

of the complex partners is strongly screened. The depth of the minima is raised with respect to the vacuum potential.[48,54] (2) The interference of the solvent shells rises barriers between local minima (see also Figure 3). Such barriers are induced by solvent molecules "bridging" host and guest molecules and have been reported earlier for ion pairs[17,18,55,56] and the system K^+-18-crown-6 in aqueous solution.[48] Finally, Figure 6 shows the PMF along the MEP to get a clearer picture of the energetics.

It is interesting to note that the one-dimensional PMF (cf. Figure 4) is very similar to the MEP, more so for the lighter ions. We might conclude that the entropic contribution for movement orthogonal to the MEP is small and increases with the weight of the approaching ion.

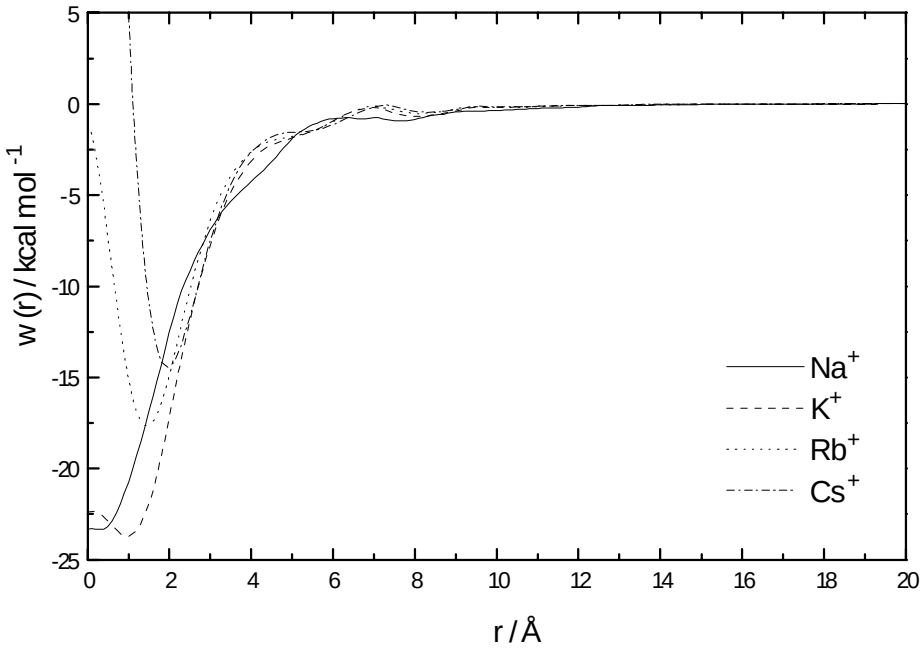

Figure 6. FES of host-guest interaction along the minimum energy path found in acetonitrile

2.5 Conclusion

We have presented a hybrid RISM/MC technique specifically suited for rapid determination of free energy surfaces for bimolecular association in solution and

applied it to study the complexation of alkali metal ions and 18-crown-6 in methanol and acetonitrile. The solute-solute configurational space can be explored with respect to various observables: Local minimization subject to the a fixed distance constraint leads to the identification of minimum energy paths that are associated with the reaction mechanism in solution. Related information is gained from two-dimensional slices through the FES. By using the free energy perturbation method during sampling of the solute degrees of freedom it is possible to construct a radially averaged potential of mean force from which the computation of the free energy of complexation is possible. The main computational advantage of our method stems from the existence of an incremental formula that relates a change of chemical potential to a change of conformation starting from one initial set of RISM solutions. It is therefore not necessary to recompute the explicit RISM solution for every new configuration, thus accelerating the computations substantially.

Interesting features of the FES were revealed by determination of minimum energy paths: Pathways with three-fold symmetry leading to the global minimum were found that favor a lateral attack of the ions. Furthermore, a solvent-induced barrier shields a collinear attack. A similar structure of the FES of the system potassium-18-crown-6 in water has been proposed earlier. Although the 1D-RISM equations within the HNC approximation suffer from well-known limitations, the results obtained in this study agree reasonably well with experiment, notably better in the case of methanol. Within each series, the selectivity of the crown is found to be consistent with the experimental data. Although the D_{3h}-symmetric geometry is the most stable conformation of 18-crown-6 in polar solvents, the rigidity of the model probably accounts for most of the differences between the calculated and the experimentally measured free energies of complexation. Furthermore, the simplicity of the solvent models and the neglect of counterions may pose problems.[57] We might conclude that 1D-RISM/HNC theory is well suited at least for the prediction of relative effects for similar complexes in solution. The transfer to other solvents requires a careful check of the models involved.

Ongoing work is concerned with the prediction of chiral selectivity in solution. The methods presented here will most likely be suited for this task that would need huge computational power if done with molecular simulation using explicit solvent models. For increasingly complex systems, more elaborate techniques of minimization on the FES will be needed.

2.6 References

(1) Reddy, M. R.; Erion, M. D.; Agarval, A. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; Wiley-VCH: New York, 2000; Vol. 16, p 217.

(2) Roux, B.; Simonson, T. *Biophys. Chem*. **1999**, *78*, 1.

(3) Hirata, F.; Rossky, P. J. *Chem. Phys. Lett.* **1981**, *83*, 329.

(4) Enciso, E. *Mol. Phys.* **1985**, *56*, 129.

(5) Kovalenko, A.; Ten-no, S.; Hirata, F. *J. Comp. Chem.* **1999**, *20*, 928.

(6) Chiles, R. A.; Rossky, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 6867.

(7) Pettitt, M. B.; Karplus, M. *Chem. Phys. Lett.* **1987**, *136*, 383.

(8) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 1712.

(9) Lee, P. H.; Maggiora, G. M. *J. Phys. Chem.* **1993**, *97*, 10175.

(10) Svensson, B.; Woodward, C. E. *J. Phys. Chem.* **1995**, *99*, 1614.

(11) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *J. Phys. Chem.* **1996**, *100*, 1111.

(12) Melenkevitz, J.; Schweitzer, K. S.; Curro, J. G. *Macromolecules* **1993**, *26*, 6190.

(13) Khalatur, P. G.; Khokhlov, A. R. *Mol. Phys.* **1998**, *93*, 555.

(14) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Am. Chem. Soc.* **1998**, *120*, 1855.

(15) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Chem. Phys.* **1999**, *110*, 4090.

(16) Pettitt, M. B.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 5836.

(17) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10391.

(18) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10403.

(19) Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **1999**, *103*, 7942.

(20) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *113*, 2793.

(21) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 1930.

(22) Hansen, J. P..; McDonald, I. R. *Theory of simple liquids*, Academic Press: London **1980**.

(23) Eu, B. C.; Rah, K. *J. Chem. Phys*. **1999**, *111*, 3327.

(24) von Solms, N.; Chiew, Y. C. *J. Chem. Phys.* **1999**, *111*, 4839.

(25) Fries, P. H.; Patey, G. N. *J. Chem. Phys.* **1985**, *82*, 429.

(26) Rossky, P. J.; Pettitt, B. M.; Stell, G. *Mol. Phys.* **1983**, *50*, 1263.

(27) Singer, S. J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

(28) Pettitt, B. M.; Karplus, M.; Rossky, P. J. *J. Phys. Chem.* **1986**, *90*, 6335.

(29) Lau, W. F.; Pettitt, B. M. *Biopolymers* **1987**, *26*, 1817.

(30) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett.* **1985**, *121*, 194.

(31) Karplus, M. CHARMM Documentation, The CHARMM Development Project, Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

(32) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.

(33) Prue, J. E. *J. Chem. Educ.* **1969**, *46*, 12.

(34) Pranata, J.; Jorgensen, W. L. *Tetrahedron* **1991**, *47*, 2491.

(35) Dang, L. X.; Kollman; P. A. *J. Am. Chem. Soc.* **1990**, *112*, 5716.

(36) Steinberg, I. Z.; Scheraga, H. A. *J. Biol. Chem.* **1963**, *238*, 172.

(37) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047.

(38) Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707.

(39) Luo, R.; Head, M. S.; Given, J. A.; Gilson, M. K. *Biophys. Chem.* **1999**, *78*, 183.

(40) Hill, T.L. *Statistical Thermodynamics*, Addison-Wesley Publishing: London **1950**.

(41) Dang, L. X.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 55.

(42) Jorgensen, W. L. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(43) Jorgensen, W. L. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(44) Wipff, G.; Weiner, P.; Kollman, P. A. *J. Am. Chem. Soc.* **1982**, *104*, 3249.

(45) Talman, J. D. *J. Comput. Phys.* **1978**, *29*, 35.

(46) Rossky, P. J.; Friedman, H. L. *J. Chem. Phys.* **1980**, *72*, 5694.

(47) Jorgensen, W. L. *User's manual of BOSS, Version 4.1* **1999**, Department of Chemistry, Yale University, New Haven, CT.

(48) Kowall, T.; Geiger, A. *J. Phys. Chem.* **1995**, *99*, 5240.

(49) Izatt, R. M.; Bradshaw, J. S.; Nielsen, S. A.; Lamb, J. D.; Christensen, J. J. *Chem. Rev.* **1985**, *85*, 271.

(50) Izatt, R. M.; Pawlak, K.; Bradshaw, J. S.; Bruening, R. L. *Chem. Rev.* **1995**, *95*, 2529.

(51) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vettering, W. T. *Numerical Recipes in C*, Cambridge University Press: Cambridge 1988.

(52) Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C. D. In *Data Visualization in Molecular Science - Tools for Insight and Innovation*; Bowie, J. E.; Ed.; Addison-Wesley: Reading 1995.

(53) Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer III, H.F.; Schreiner, P.R.; Eds.; Wiley: Chichester 1998.

(54) Wipff, G.; Weiner, P.; Kollman, P. A. *J. Am. Chem. Soc.* **1982**, *104*, 3249.

(55) Dang, L. X.; Pettitt, B. M. *J. Phys. Chem.* **1990**, *94*, 4303.

(56) Gao, J.; Boudon, S.; Wipff, G. *J. Am. Chem. Soc.* **1991**, *113*, 9610.

(57) Dang, X. D. *Chem. Phys. Lett.* **1994**, *227*, 211.

3 Theoretical study of hydrogen-bonded cation-crown ether complexes with a hybrid integral equation/Monte-Carlo approach

3.1 Summary

The hybrid reference interaction site model/Monte-Carlo (RISM/MC) approach to complexation free energy surfaces (FES) in solution has been extended to the interaction of molecular ions and uncharged species. Therein, a fast integral equation model of the solvent phase is used to derive thermodynamic and structural properties of molecular solvation.

Monte Carlo (MC) simulation techniques are employed to sample the rigid body solute-solute interaction FES effectively. Interaction potentials of mean force are presented for the binding of organic ammonium cations and some guanidinium derivatives to 18-crown-6 and 15-crown-5 in water and methanol. By radial integration of the potentials of mean force we calculate complexation free energies of binding which are further corrected with respect to standard state conditions of 1 mol l^{-1} and 298.15 K.

The initial structures of the complexes are prepared by energetic minimization of the solid state structures with a fully flexible force field representation. By this means, the TRIPOS all atom force field was used in conjunction with an integral equation representation of the solvent and a simplex downhill algorithm. It is shown that the 1D integral equation formulation of the solvent is capable to provide a reasonable representation of the solvent distribution and around a solute molecule. The calculated standard free energies of binding are compared with the experimental data.

3.2 Introduction

The rapid development of supramolecular chemistry has attracted much interest in the computational treatment of host-guest interaction. A remarkable

amount of experimental data exists regarding the complexation of macrocyclic polyethers with cationic, anionic and even neutral guests under various solvent conditions.[1,2,3,4,5,6,7] Therein, the binding mechanisms for organic guest molecules differ from those of monovalent metal ions. Metal cations form inclusion complexes with the macrocyclic rings while organic ammonium derivatives hydrogen bond to certain donor atoms.[8,9,10] Often a high degree of selectivity for a certain sort of guest molecules is found with crown-ethers, e.g. 18-crown-6. From a phenomenological point of view, these findings are readily explained with the molecular size, sterical hindrance or conformational flexibility of host-guest interaction. In contrast, thermodynamic studies offer a deeper understanding of the driving forces and structure-function relationship, that lead to specific complexation in solution. Among other factors, the solvent plays a very important role for the stability and selectivity of the crown-ether complexes. Polar solvent molecules compete with guest molecules for bonding positions since both host and guest molecules loose parts of their solvation shells during the complexation process.

Different computational approaches have been applied to study host-guest complexation in solution, e.g. quantum-chemical methods,[11,12] molecular dynamics calculations,[13,14,15,16] Monte-Carlo simulations,[17,18,19] integral equation techniques[20,21,22,23] and empirical approaches,[24,25] which often lack an appropriate representation of the solvent. While simulation techniques are still the most popular approaches to thermodynamics of solutions, integral equation techniques have recently gained some interest in the literature.[26,27,20-23,28,29,30] Integral equation techniques benefit from a statistical model of the solvent which is derived quickly in comparison to the large computational effort commonly spent in the sampling of solvent molecules. In the present work, we apply a hybrid reference interaction site model integral equation/Monte Carlo approach (RISM/MC)[26] to the determination of supramolecular complexation free energy surfaces (FES) in solution. This method combines a fast statistical solvent model[27] and Monte-Carlo sampling of solute-solute interaction[17-19,28-30] with the technique of statistical perturbation.[15,18,31] Our past approach has been extended to describe the interaction of molecular ions and uncharged species in various solvents.

We report orientationally averaged potentials of mean force for the anchoring of substituted ammonium cations and some guanidinium derivatives (fig 1) with

15-crown-5 and 18-crown-6 in water and methanol. Free energies of complexation are extracted from the PMFs and compared with experimental data.[1-10]

This paper is outlined as follows. In the next section the hybrid Monte Carlo / integral equation approach is described. Section 3 provides computational details of the methods used. Structural and thermodynamic data are reported in section 4 and summarized in the final section.

3.3 Methods

3.3.1 Integral equation theory

Our calculations are performed using the reference interaction site model (RISM) originating from Hirata and Rossky[32] and Chandler and Andersen.[33] The RISM matrix integral equations for a pure solvent of density ρ relates the intermolecular pair correlation function matrices **h** as functionals of the intramolecular correlation functions **ω** and the direct correlation functions **c**:

$$\mathbf{h}^{vv} = \boldsymbol{\omega}^{v} * \mathbf{c}^{vv} * \boldsymbol{\omega}^{v} + \boldsymbol{\omega}^{v} * \mathbf{c}^{vv} * \rho \mathbf{h}^{vv} \tag{3.1}$$

All matrices are site-site labeled v · v matrices, where v is the number of solvent interaction sites. The star (*) is a short notation representing matrix convolution products. The intramolecular correlation matrix in real space representation for a rigid molecule consists of delta functions which constrain the atomic interaction sites *i*, *j* to the bondlength $\mathbf{d}_{ij}$:

$$\omega_{ij}(r) = (1-\delta_{ij})\frac{\delta(|\mathbf{r}| - |\mathbf{d}_{ij}|)}{4\pi |\mathbf{d}_{ij}|^2} \tag{3.2}$$

In the infinite dilution limit, the properties of the bulk phase remain unchanged under insertion of a solute, and the solvent-solvent susceptibility matrix $\boldsymbol{\chi}^{vv}$ is written as

$$\boldsymbol{\chi}^{vv} = \boldsymbol{\omega}^{v} + \rho \mathbf{h}^{vv} \tag{3.3}$$

The integral equation for the solute-solvent correlation functions $\mathbf{h}^{Xv}$ then reads

$$\mathbf{h}^{Xv} = \boldsymbol{\omega}^{X} * \mathbf{c}^{Xv} * \boldsymbol{\chi}^{vv} \tag{3.4}$$

The dimensionalities of $\boldsymbol{\omega}^{X}$ and $\mathbf{h}^{Xv}$ are X · X and X · v, respectively. The HNC closure[27] for the matrix elements *ij* is written as

$$h_{ij} = \exp\left(-\frac{U_{ij}}{kT} + h_{ij} - c_{ij} \right) - 1 \tag{3.5}$$

where k is the Boltzmann constant and T is the absolute temperature. U_{ij} is the pair potential between the interaction sites ij, which is modeled in accordance to standard force fields by a short-ranged Lennard-Jones part and a long-ranged Coulomb interaction:

$$U_{ij} = U_{ij}^{\mathrm{LJ}} + U_{ij}^{\mathrm{C}} \tag{3.6}$$

A detailed description of the force field used in this study is given in the following section.

The correlation function matrices $\mathbf{h}^{\mathrm{Xv}}$ and $\mathbf{c}^{\mathrm{Xv}}$ are calculated by solving the set of coupled integral equations 3.1, 3.4 and 3.5. Free energies of solvation $\Delta\mu$ for a solute in a given conformation $\boldsymbol{\omega}^{\mathrm{X}}$ are derived by radial integration of the correlation matrices.[34]

$$\Delta\mu\left(\boldsymbol{\omega}^{\mathrm{X}}\right) = -\frac{\rho}{\beta} \sum_{i,j} \int d\mathbf{r} \left(\frac{1}{2} h_{ij}^2 - c_{ij} + \frac{1}{2} h_{ij} c_{ij} \right) \tag{3.7}$$

One has, in principle, to solve the set of RISM equations fully to obtain free energies of solvation for any solute conformation $\boldsymbol{\omega}^{\mathrm{X}} \neq \boldsymbol{\omega}^{\mathrm{Y}}$. However, the difference in the free energy of solvation related to a structural change (e.g. a torsion) is given by:[26,35]

$$\Delta\Delta\mu\left(\boldsymbol{\omega}^{\mathrm{X}}, \boldsymbol{\omega}^{\mathrm{Y}}\right) = -\frac{\rho\, kT}{2} \sum_{i,i'} \int d\mathbf{r} \left[\mathbf{c}^{\mathrm{Xv}} * \boldsymbol{\chi}^{\mathrm{vv}} * \mathbf{c}^{\mathrm{Yv}} \left(\boldsymbol{\omega}^{\mathrm{X}} - \boldsymbol{\omega}^{\mathrm{Y}} \right) \right]_{i,i'} \tag{3.8}$$

The summation has to performed over all sites $\{i;i'\}$ of the solutes X and Y. No restrictions are applied to the nature of the solute X and the perturbed structure Y, which might be a single molecule, a constrained supramolecule or even a set of molecules. The full solution of the RISM equations 3.1, 3.4 and 3.5 is needed for a single configuration only, while the relative free energy of other configurations is given by the convolution equation 3.8. This opens a computationally simple route to include solvation thermodynamics in local optimization of intramolecular degrees of freedom as well as the sampling of solute-solute interaction.

3.3.2 Local optimization

Whenever a crystal structure is available for a complex, it offers a reasonable starting geometry for docking studies in solution. However, complex structures in solution can deviate from the solid state crystal structure. But, an approximation of the solvation geometry is readily calculated by an optimization procedure with a classical vacuum force field. Accounting for site-site solute-solvent interaction, the RISM model allows for implementation of solvent terms in optimization procedures. We therefore estimated the solution structure of the anchored components in solution by a local intramolecular and intermolecular relaxation including the surrounding RISM solvent in a single step process. To this end, the TRIPOS intramolecular force field[36,37] was extended by a solvent contribution from eq 3.8 and combined with a simplex algorithm.[38] The vacuum energy ΔV_{XY} then reads

$$\begin{aligned}\Delta V_{XY} = &\sum_{i}^{\text{bonds}} \frac{1}{2} k_i^{\text{bond}} \left(d_i - d_i^{\text{o}}\right)^2 + \sum_{j}^{\text{angles}} \frac{1}{2} k_j^{\text{angle}} \left(\theta_j - \theta_j^{\text{o}}\right)^2 \\ &+ \sum_{k}^{\text{torsions}} \frac{1}{2} V_k^{\text{torsion}} \left[1 + S_k \cos\left(|n_k| \cdot \vartheta_k\right)\right] \\ &+ \sum_{l}^{\text{out-of-plane}} \frac{1}{2} k_l^{\text{oop}} {d_l}^2 + E_{\text{nonbonded}}\end{aligned} \tag{3.9}$$

where k_{i}, k_{j}, k_{l} represent the bond stretching, angle bending and out-of-plane bending constants, respectively. Equilibrium values for the bondlengths and the angles are termed d_{i}^{o} and $\theta_{\text{j}}^{\text{o}}$. V_{k} is the torsional barrier of the k^{th} torsion, n_{k} is its periodicity and ϑ_{k} is the torsion angle. S_{k} equals to one, if the minimum energy of the torsion is found in staggered conformation, and equals to -1, if the eclipsed conformation is preferred. The total energy E_{tot} is given by

$$E_{\text{tot}}\left(\boldsymbol{\omega}^{\text{XY}}\right) = \Delta V_{\text{XY}}\left(\boldsymbol{\omega}^{\text{XY}}\right) + \Delta\mu\left(\boldsymbol{\omega}^{\text{XY}}\right) \tag{3.10}$$

This algorithm opens the possibility to estimate solvent effects on the structure of molecular ensembles quickly.

A robust simplex downhill algorithm[38] is employed to perform local optimization of complexes in a RISM solvent. A simplex is the geometrical figure in n dimensions consisting of n+1 vertices. The simplex algorithm takes such a set of n+1 points and uses reflection, expansion and contraction steps to move them into a local minimum. Taking the initial guess as $\mathbf{r}_{\text{o}}$, the simplex is initialized as

$$\mathbf{r}_n = \mathbf{r}_o + \lambda_n \mathbf{e}_n \tag{3.11}$$

where the **e** are unit vectors and λ is 0.1 (Å) for cartesian optimizations and 5 (degrees) for angular coordinates. For a given supramolecular complex XY, eq 3.10 is minimized until the convergence criterion

$$E_{tot}\left(\boldsymbol{\omega}^{XY}\right) - E_{tot}\left(\boldsymbol{\omega}^{XY'}\right) < 0.0001 \text{ kcal mol}^{-1} \tag{3.12}$$

is met.

3.3.3 Construction of free energy surfaces

Eq 3.8 has been successfully applied in the calculation of the thermodynamics of alkali ion-crown ether complexes.[26] Therein, molecular flexibility was not considered at all, since under polar solvent conditions the ring of 18-crown-6 is known to be organized in a D_{3h} symmetric conformation.[15,39,40] Eq 3.8 was employed to calculate solvation free energy differences due to solute-solute *intermolecular* degrees of freedom.

The total free energy of interaction $W(\mathbf{\Gamma})$between nonspherical solute molecules depends on six intermolecular coordinates $\mathbf{\Gamma} = \{r, \theta, \varphi, \Theta, \Phi, \Psi\}$,

$$W(\mathbf{\Gamma}) = -\Delta\mu_X - \Delta\mu_Y + \Delta V_{XY}(\mathbf{\Gamma}) + \Delta\mu_{XY}(\mathbf{\Gamma}) \tag{3.13}$$

where $\Delta\mu_X$ and $\Delta\mu_Y$ are the free energies of solvation of molecules X and Y, $\Delta\mu_{XY}(\mathbf{\Gamma})$ is the free energy of solvation of a constrained supermolecule XY and $\Delta V_{XY}(\mathbf{\Gamma})$ is the vacuum pair potential. Eq 3.13 defines the free energy surface for rigid-body interaction of two complex partners, which is sampled by variation of the intermolecular coordinates. In principle, eqs 3.8 and 3.13 allow to calculate energetics with respect to arbitrary intermolecular degrees of freedom. But, the inclusion of a fully flexible intramolecular force field increases the numerical costs dramatically.

3.3.4 Sampling of free energy surfaces

The details of the Monte-Carlo procedure used to compute orientationally averaged potentials of mean force from $W(\mathbf{\Gamma})$ have been described in a preceding publication.[26] The reaction coordinate r was chosen as the distance between the center-of mass of the crown ether and the nitrogen atom of the guest molecule. Using

statistical perturbation theory,[15,18,31] the differential potential of mean force (PMF) $w(r)$ for a small perturbation of the reaction coordinate r is:

$$w(r') - w(r) = -kT \ln\left\langle \exp\left(\frac{W(\mathbf{\Gamma}) - W(\mathbf{\Gamma}')}{kT} \right) \right\rangle_r \tag{3.14}$$

r and r' denote the reference, and perturbed reaction coordinates, respectively. Solute-solute statistics are calculated in a series of Monte-Carlo steps at 298.15 K. The resulting PMF contains the information about the free energy surface which is essential to calculate free energies of binding of the complex partners. The PMF is related to the free energy of binding of a host-guest complex:[15,18,31,41]

$$\Delta G_{\mathrm{XY}} = -k_{\mathrm{B}} T \ln\left\{ \frac{4\pi}{V_{\mathrm{o}}} \int_0^{r_{\mathrm{c}}} \mathrm{r}^2 \exp\left(-w(r)/k_{\mathrm{B}}T\right) \mathrm{dr} \right\} \tag{3.15}$$

The borders of the integration interval $[0;r_{\mathrm{c}}]$ are given by the binding region, e.g. from zero to the dissociation limit where the PMF becomes positive.

Even if the simulations are taken at the infinite dilution limit, eq 3.15 provides free energies of binding related to a standard concentration of one molecule Å^{-3}. Using the results of Gilson et al.[41], we apply a correction for the standard state (1 mol l^{-1}):

$$\Delta G^{\mathrm{o}}_{\mathrm{XY}} = \Delta G_{\mathrm{XY}} + RT \ln \frac{8\pi^2 \cdot s_{\mathrm{XY}}}{s_{\mathrm{X}} \cdot s_{\mathrm{Y}}} - RT \ln C^{\mathrm{o}} V_{\mathrm{o}} \tag{3.16}$$

where C^{o} is 1660 Å^{-3} and the volume term V^{o} equals to 1 Å^3. s_{XY}, s_{X} and s_{Y} are the rotational symmetry numbers according to Hill.[42]

3.4 Computational details

All nonbonded parameters were adopted from the *optimized potential for liquid simulation* (OPLS) force field of Jorgensen and coworkers.[43] Whenever necessary, scaling of the atomic partial charges was performed to preserve the molecular charges.

$$U_{ij} = U_{ij}^{\mathrm{LJ}} + U_{ij}^{\mathrm{C}} = 4\pi\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \kappa \frac{q_i q_j}{r_{ij}} \tag{3.17}$$

Intermolecular Lennard-Jones parameters were calculated according to the Lorenz-Berthelot combination rules:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \tag{3.18}$$

To take account for accurate screening of the long range interaction,[44] the RISM solvent-solvent intermolecular Coulombic potential is replaced by a phenomenological result:

$$\kappa = \begin{cases} \left[\dfrac{(\varepsilon_r - 1) k_B T + \frac{4}{3} \varepsilon_o \pi \rho p^2}{(\varepsilon_r - 1) \cdot \frac{4}{3} \pi \rho p^2} \right] & ; \text{solvent} - \text{solvent interaction} \\ 1 & ; \text{solute} - \text{solvent interaction} \end{cases} \tag{3.19}$$

p is the absolute value of the dipole moment of the solvent molecule, ε_o is the vacuum dielectric constant and ε_r is the relative dielectric constant of the solvent. The dielectric constants ε_r of water and methanol were set as 78.3 and 32.66, respectively. Three-site models for TIP3P water and methanol were taken from the OPLS force field.[45] However, to avoid coulombic singularities of non-repulsive interaction sites, all Lennard-Jones parameters of polar hydrogens were set to $\sigma_H = 0.4$ Å and $\varepsilon_H = 0.046$ kcal mol^{-1}. The nonbonded parameters of 15-crown-5 and 18-crown-6 were also taken from the OPLS force field. Available crystal structures of 15-crown-5 and 18-crown-6 complexes were adopted from the Cambridge Structural Data base (CSD).[46] The guest molecules were modeled by use of the SYBYL program package[36] and positioned at the binding site. This step was followed by a relaxation of all coordinates in presence of a RISM solvent (eq 3.10). The resulting structures were used as input data for the RISM/MC algorithm.

The HNC-RISM equations were solved on a logarithmically spaced grid of 512 points, ranging from a minimum of $5.98 \cdot 10^{-3}$ Å to a maximum of 164.02 Å. For the calculation of forward and backward fast Fourier transforms (FFT) we used the Talman procedure.[47,48] Number densities of liquid water, methanol and acetonitrile at 298.15 K were taken as 0.03334 Å^{-1}, 0.01405 Å^{-1} and 0.0116 Å^{-1}, respectively. The solvent-solvent and solute-solvent integral equations were iteratively converged until the convergence criterion for the direct correlation function $\max(\Delta c_{ij}) < 10^{-5}$ was met.

TABLE 1: Force field parameters: atomic partial charges *q*; Lennard-Jones parameters σ and ε

Site	*q* / e	σ / Å^{-1}	ε / kcal mole^{-1}	Site	*q* / e	σ / Å^{-1}	ε / kcal mole^{-1}
Water				**A_4**			
O	-0.8340	3.1508	0.1521	C	0.1420	3.7500	0.1050
H	0.4170	0.4000	0.0460	O	-0.3900	2.9600	0.2100
Methanol				N	-0.5420	3.2500	0.1700
O	-0.7000	3.070	0.1700	H	0.3330	0.5000	0.0300
CH_3	0.2650	3.775	0.1180	**A_5**			
H	0.4350	0.400	0.0460	C	0.6400	2.2500	0.0500
15-crown-5				N	-0.8000	3.2500	0.1700
C	0.1400	3.500	0.0660	H	0.4600	0.5000	0.0300
O	-0.4000	2.900	0.1400	**A_6**			
H	0.0300	2.500	0.0300	C	0.6400	2.2500	0.0500
18-crown-6				$N(H_2)$	-0.8000	3.2500	0.1700
C	0.1400	3.500	0.0660	H(N)	0.4600	0.5000	0.0300
O	-0.4000	2.900	0.1400	$N(CCH_2)$	-0.7000	3.2500	0.1700
H	0.0300	2.500	0.0300	C(H)	0.2000	3.5000	0.0660
A_1				H(C)	0.0700	2.5000	0.0150
N	-0.4000	3.2500	0.1700	**A_7**			
H	0.3500	0.5000	0.0300	C(N)	0.6400	2.2500	0.0500
A_2				N	-0.8000	3.2500	0.1700
N	-0.3000	3.2500	0.1700	H(N)	0.4600	0.5000	0.0300
C	0.1300	3.5000	0.0660	$C(CN_2H_4)$	0.1200	3.5500	0.0700
H(C)	0.0600	2.5000	0.0300	C(H)	-0.1150	3.5500	0.0700
H(N)	0.3300	0.5000	0.0300	H(C)	0.1150	2.4200	0.0300
A_3							
N	-0.3000	3.2500	0.1700				
H(N)	0.3300	0.5000	0.0300				
$C(H_2)$	0.1900	3.5000	0.0660				
$C(H_3)$	-0.1800	3.5000	0.0660				
H(C)	0.0600	2.5000	0.0300				

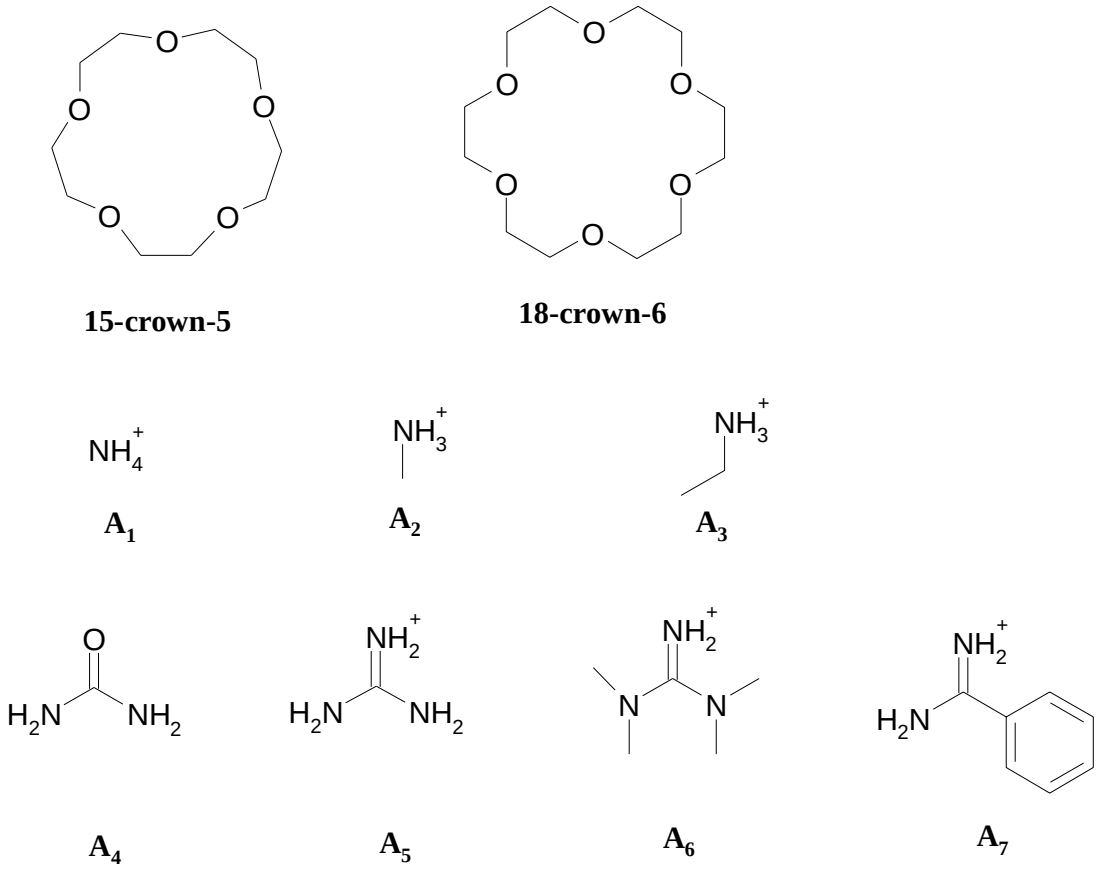

Figure 1: Molecular structures of two crown ether host molecules 15-crown-5 and 18-crown-6, three primary ammonium guests A_1, A_2 and A_3, and four guanidinium-type guests urea (A_4), guanidine (A_5), 1,1,3,3-tetramethylguanidine (A_6) and benzamidine (A_7) whose complexation thermodynamics have been studied with the RISM/MC technique.

3.5 Results and Discussion

3.5.1 Solvation structure of 18-crown-6

The RISM equations were solved for 18-crown-6 in water and methanol. The radial distribution functions (rdf) for the interaction of the crown-ether oxygen to the solvent atoms are given in Figs. 2-5. Due to hydrogen bonding, the O_{crown}-H_{water} rdf (cf. fig 2) exhibits a strong first peak with a maximum at 1.5 Å. With the given hydrogen number density in solution of 0.06668 $Å^{-1}$ this peak integrates to 1.03. This result is in accordance to a hydration structure of 18-crown-6 where each oxygen

atom of the crown ether is bonded to water. The first maximum of the O_{crown}-$H_{methanol}$ rdf at 1.5 Å (cf. fig 4) is even stronger. This peak integrates to a total of 0.677, predicting a first solvation shell where 4 of 6 ether oxygens are hydrogen-bonded by methanol. In contrast to the O-H distribution functions, the C-H and C-O rdfs of both water (fig 3) and methanol (fig 5) exhibit shallow peaks due to the hydrophobic nature of the methylene groups of the crown ether.

The structural information available from 1D-RISM is limited due to the 1D representation of the radial distribution functions. It is more illustrative to give a graphical illustration with explicit solvent molecules included. By this means, preferential solvent configurations near the solute can be estimated by adding explicit solvent molecules and searching the configurational space for minimum energy conformers. By this means, the additional solvent molecules are formally treated as complex partners of the host, 18-crown-6. To this end, single water or methanol molecules were docked to 18-crown-6 using the program package SYBYL[36] and the total energies of the resulting solvated supramolecules were minimized with eq 3.10.

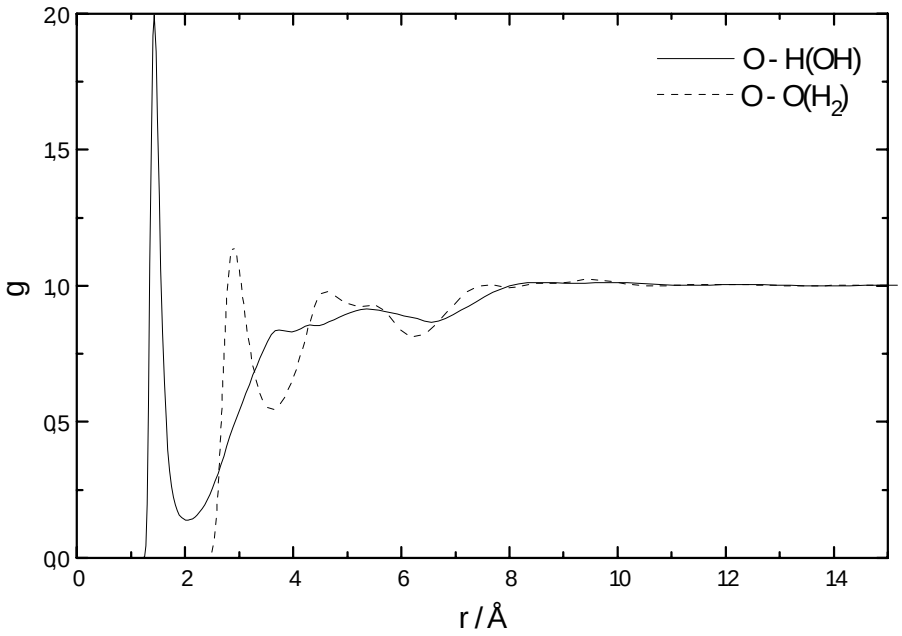

Figure 2: Radial site-site pair distribution functions for the ether oxygens of 18-crown-6 with water. The first peak of the O-H function integrates to a coordination number of 1.03.

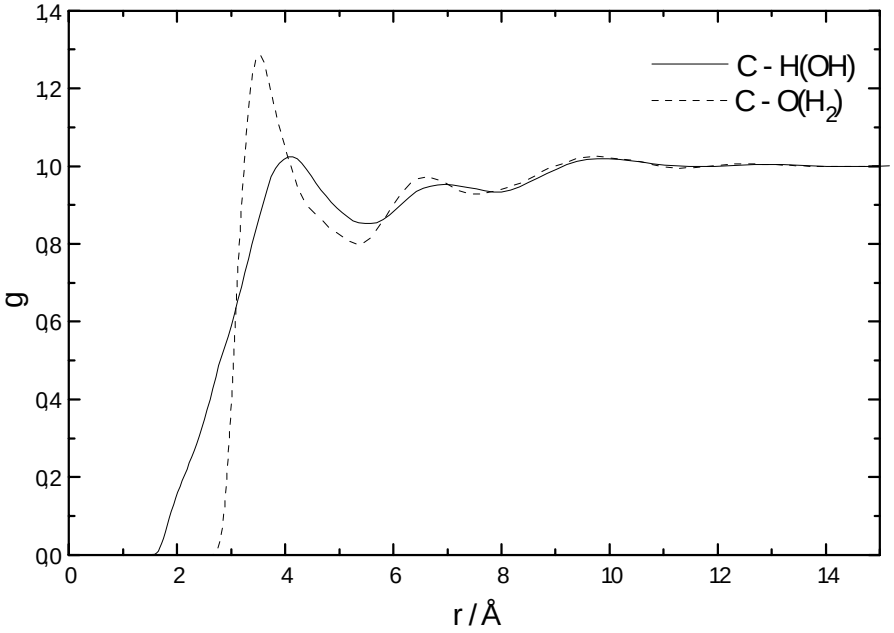

Figure 3: Radial site-site pair distribution functions for the hydrophobic ether carbons of 18-crown-6 with water, calculated by HNC-RISM.

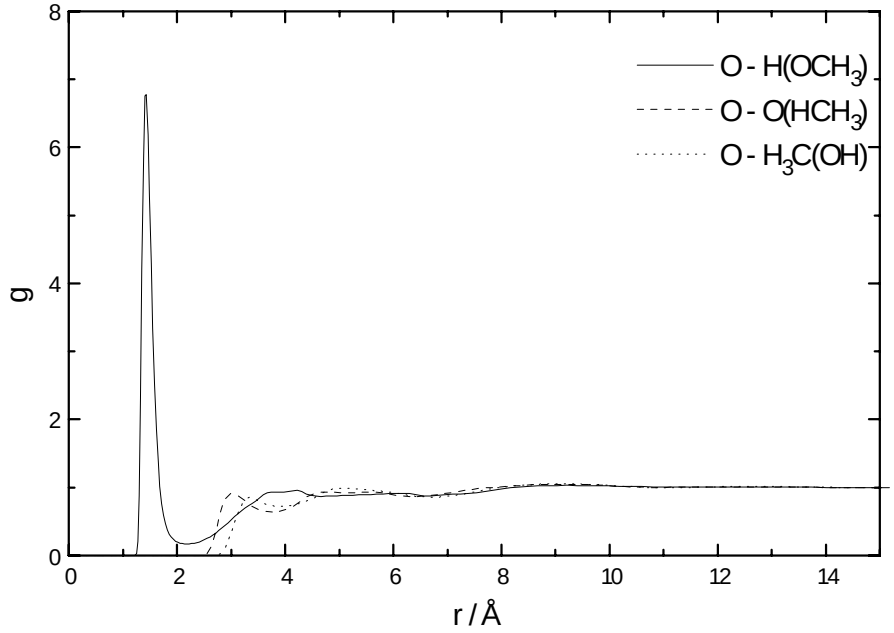

Figure 4: Radial site-site pair distribution functions for the ether oxygens of 18-crown-6 with methanol as provided by the OPLS united-atom approximation. The first peak of the O-H function integrates to a coordination number of 0.677.

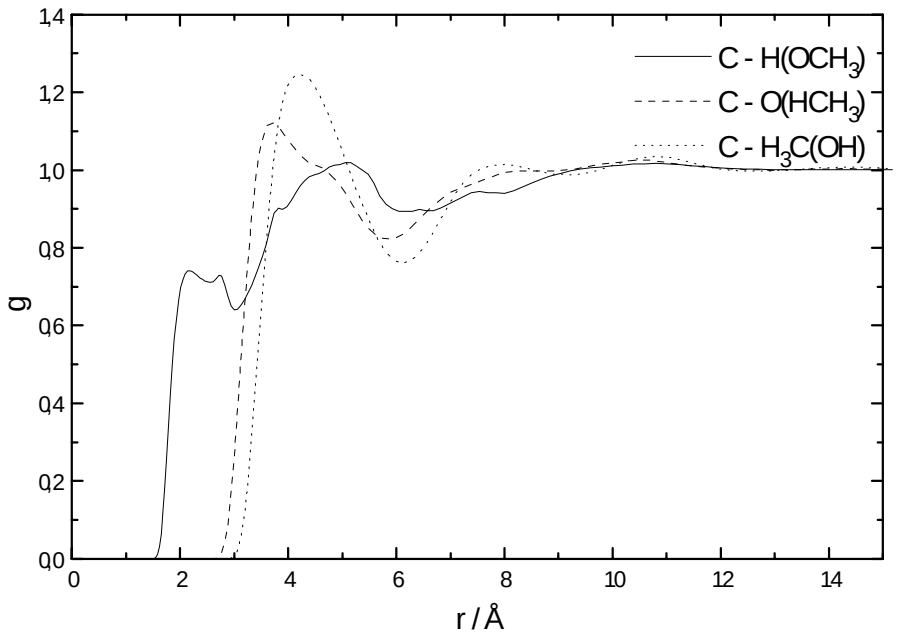

Figure 5: Radial site-site pair distribution functions for the ether carbons of 18-crown-6 with methanol, calculated by HNC-RISM.

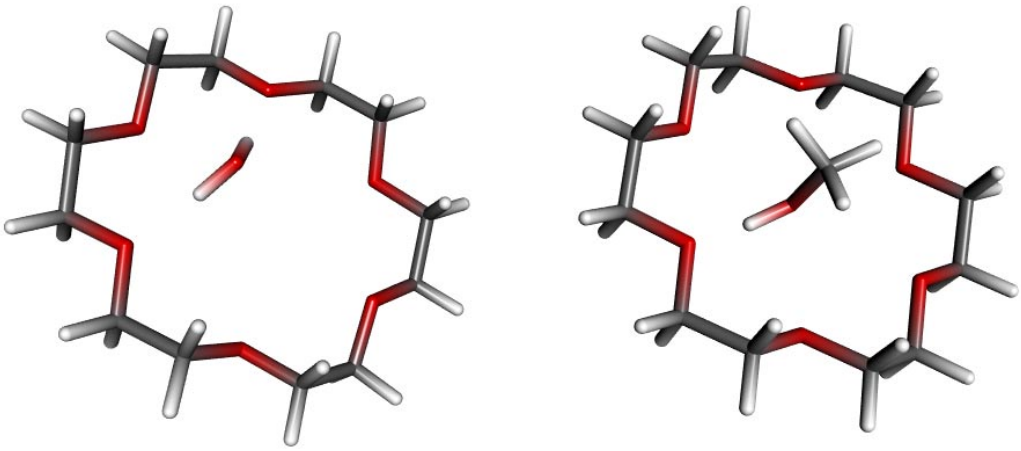

Figure 6. Minimum energy structures of 18-crown-6 complexed with water in aqueous solution (left hand side) and 18-crown-6 complexed with methanol in methanolic solution (right hand side.

In aqueous solution, a bridged configuration of the D_{3h} symmetric crown ether is found (cf. fig 6), which is in accordance to previous experimental findings and to simulation results.[40,49] The lengths of the hydrogen bonds O_{crown}-H_{water} are found to be 1.67 Å, which is very close to the first maximum of the rdf. For methanol, a hydrogen-bonded conformation[50] is predicted, with a bondlength of 1.73 Å of the O_{crown}-$H_{methanol}$ bond. This distance is also very close to the first maximum of the O_{crown}-$H_{methanol}$ rdf. The minimum-energy conformation of 18-crown-6 derived by this method differs slightly between methanol and water. The average O-C-C-O torsion angles are found to be approximately 55° in water and 58° in methanol in contrast to 67.5° found without solvent contributions.

The collinear shielding favoring a lateral attack as observed for the interaction of 18-crown-6 with atomic ions in water[40] and methanol[26] can well be explained with the given solvent structures. The formation of a crown-ether-ion complex in water requires the breakage of several hydrogen bonds.

3.5.2 Anchoring of R-NH_3^+/18-crown-6

3.5.2.1 Minimum-Energy-Path

In polar solvents, 18-crown-6 is preorganized for binding[51,52] in a D_{3h} symmetric conformation. In fact it has been shown that constraints of the crown ether symmetry in water result in free energy differences as low as 0.6 kcal mol^{-1} (see ref.[52]). We expect therefore that essential features of the FES will not be affected by our rigid model, which represents the minimum energy components of a complex in solution.

Due to the C_3 symmetric arrangement of the NH_3^+-group, mono-substituted ammonium cations fit the D_{3h} symmetric cavity of 18-crown-6 perfectly. In the X-ray structures of $H_3CNH_3^+$/18-crown-6, the ion is found locked by three equivalent hydrogen bonds. This complex is taken as a prototype for RNH_3^+/18-crown-6 complexes. The input structures for the RISM/MC algorithm were provided by docking the complex partners with the program package SYBYL,[36] followed by a flexible optimization step with eq 3.10. The minimum energy structures of the complexes R-NH_3^+/18-crown-6 are depicted in fig 7.

We examined the rigid-body complexation FES of the system $H_3CNH_3^+$/18-crown-6 by energetic minimization of a total set of 7500 random configurations with a simplex downhill algorithm. During the relaxation of the intermolecular coordinates **Γ**, the distance of the centers-of-mass was constrained. The constraint was varied within the range [0 Å;15 Å] in steps of 0.1 Å to result minimum energy pathways of the rigid complex partners.

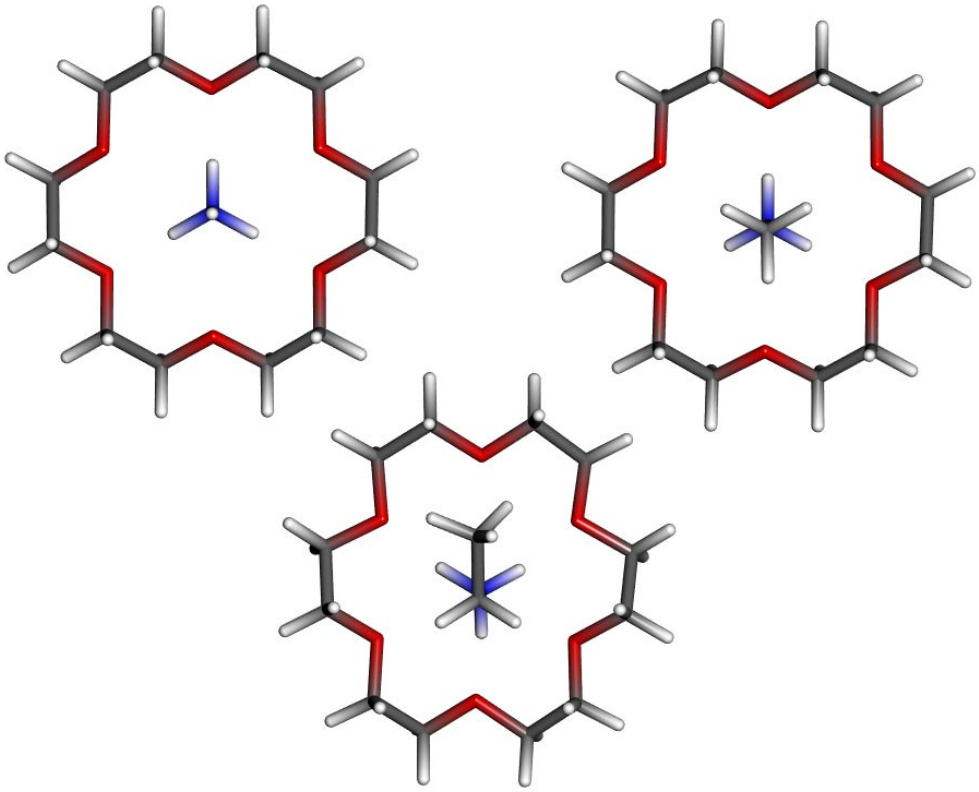

Figure 7: Minimum energy conformers of the complexes 18-crown-6/A_1/A_2/A_3 obtained by docking followed by flexible optimization in water. A perfectly symmetric network of hydrogen bonds is found. The structures obtained in methanol are visually indistinguishable.

50 configurations were minimized at each step and subsequently sorted by their relative energy. A total of 30 lowest-energy configurations were chosen at each step for visualization with the MOLCAD package[53,54] (figs 8,9). For easier interpretation of the data, we used a simple cone representation of the $H_3CNH_3^+$ ion with the bottom of each cone representing the $-NH_3^+$ residue, and the tops pointing to the positions of the corresponding carbon of the $-CH_3$ residues. The length of the

cones equals to the carbon-nitrogen bondlength. The cones are color-coded from red (bulk phase energy) to blue (most negative binding energy) to mark for the relative energies of the supramolecular assembles. The global minimum energy complex is found locked in a centered position, given by a explicit molecular representation.

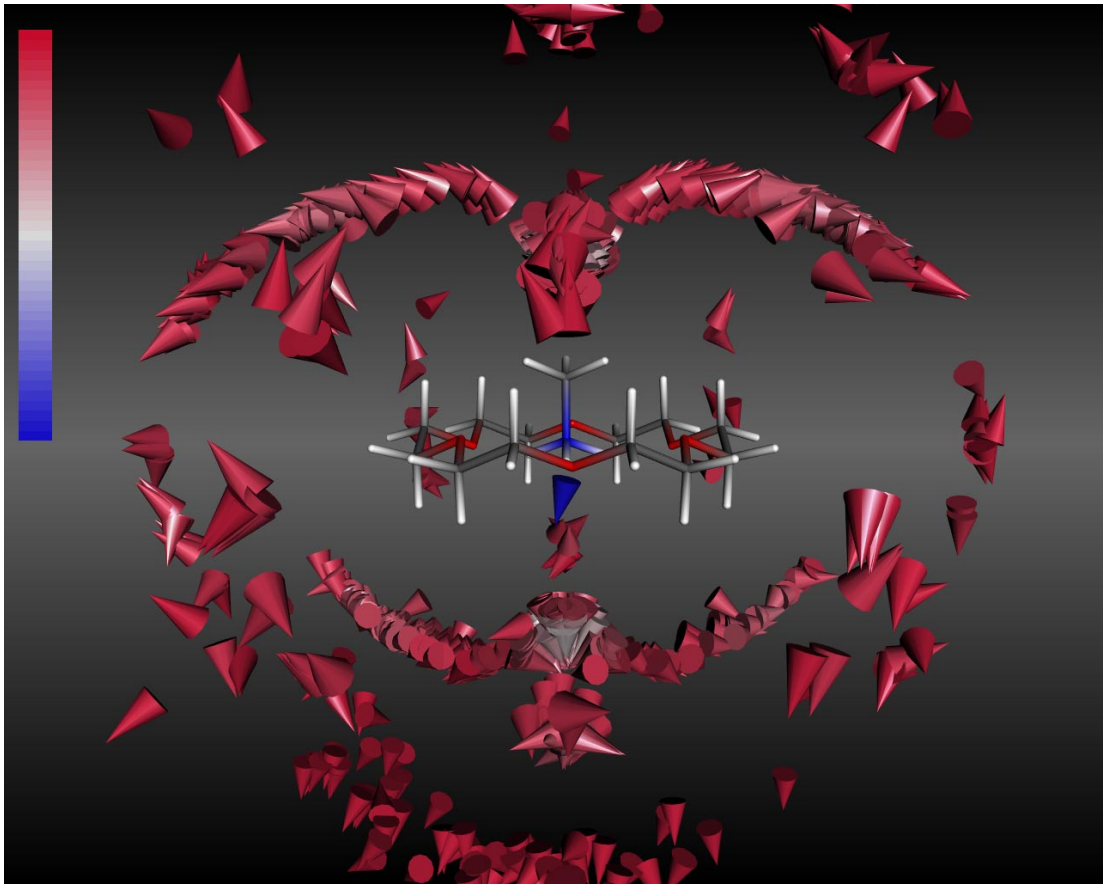

Figure 8: Minimum energy conformers of the A_2/18-crown-6 complex in water. The global minimum is depicted by a explicit molecular structure, while all other conformers are given by a cone representation. The tops of the cones give the positions of the methyl groups, and the bottoms give the positions of the corresponding amine groups. The relative energy is color-coded from red (bulk phase energy) to blue.

C_3-symmetric minimum-energy pathways are deduced from figs 8 and 9, leading to the global minimum. These pathways begin with the formation of a single hydrogen bond which allows the complex partners to rotate freely. The interaction energy drops continuously until the -CH_3 residue is located near the symmetry axis of the crown ether. The single hydrogen bond is released and three new bonds are

formed. In methanol, however, the minimum-energy pathways are longer-ranged than in water, while the preferential orientation of $H_3CNH_3^+$ is perpendicular to the molecular surface when separated from 18-crown-6 by the first solvation shell. This findings are attributed to the weaker dielectric shielding of solvent methanol in comparison to water.

At distances larger than 4 Å, clusters of possible orientations of $H_3CNH_3^+$ are found that orientates parallel to the surface of the host molecule. This is most probable due to the shielding of strong electrostatic attraction by the solvent.

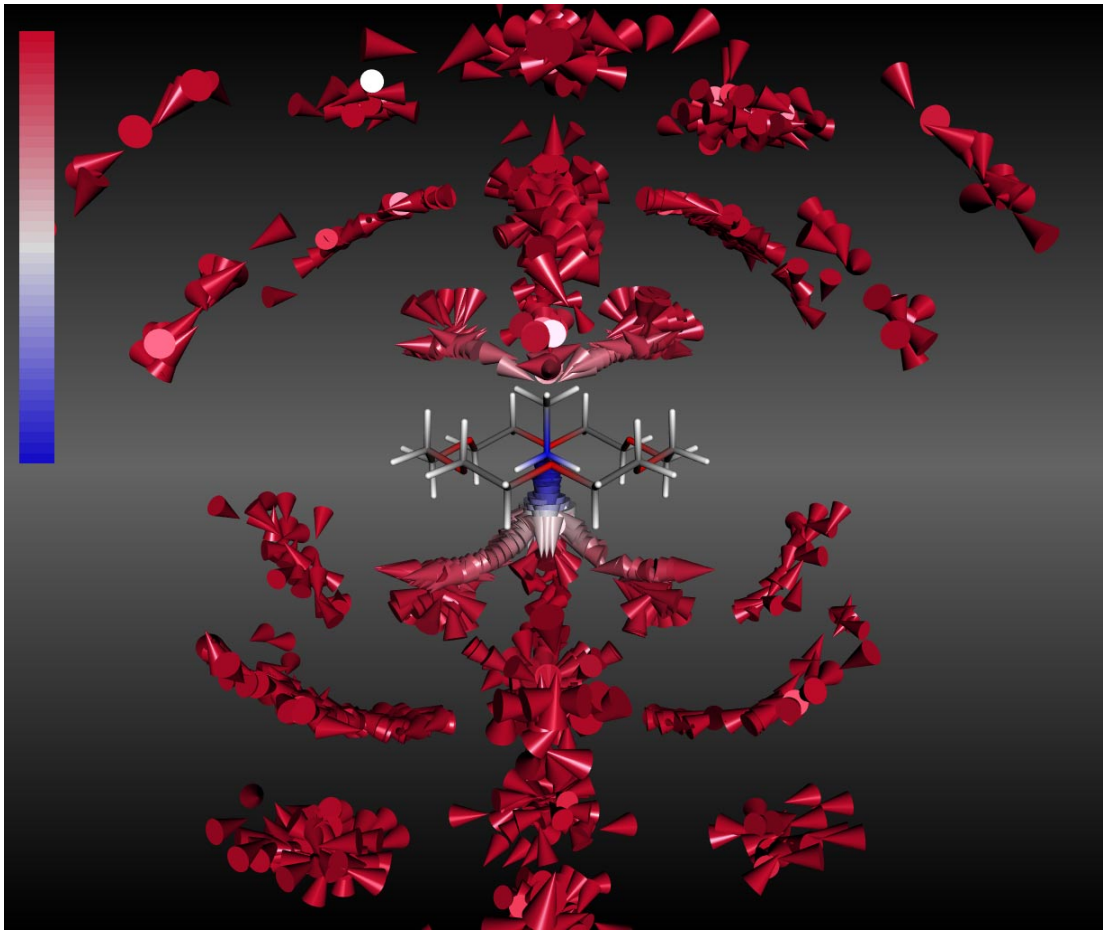

Figure 9: Minimum energy conformers of the A_2/18-crown-6 complex in methanol. The global minimum is given by a molecular representation, while all other conformers are simplified. The interaction energy is color-coded from red (bulk phase energy) to blue. Note the parallel orientation of the A_2-dipoles in the outer solvation shells.

3.5.2.2 Potentials of mean force

We have examined thermodynamic averages of the interaction of the three ammonium guests NH_4^+ (A_1), $H_3CNH_3^+$ (A_2) and $H_5C_2NH_3^+$ (A_3) with 18-crown-6. Preceding the RISM/MC calculations, the complex partners were initially placed 20 Å apart. A total of 200 perturbation windows was chosen along the reaction coordinate, each with an offset of 0.1 Å. A single perturbation window included the sampling of 20000 independent solute configurations, giving the average free energy difference for a 0.1 Å step of the guest towards the host. The corresponding potentials of mean force are given in figs 10 and 11.

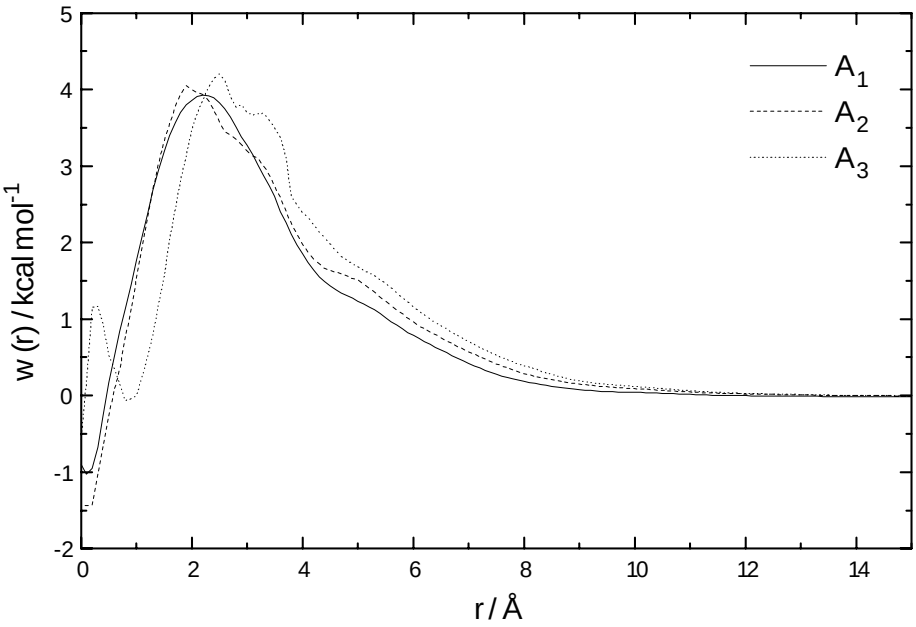

Figure 10: Potentials of mean force for the anchoring of amines A_1, A_2 and A_3 to 18-crown-6 in aqueous solution. Complexation in water comes along with releasing the strong solute-solvent interactions, therefore tiny interaction energies are found.

The data for NH_4^+ in methanol is missing since a stable solution of the RISM equations could numerically not be obtained. In both methanol and water, well-defined host-guest binding positions are found, corresponding to a close contact of host and guest molecules. Clearly, the ionic solvation in water is considerably stronger than it is in methanol. The average barrier height found in water is ~4 kcal mol^{-1} compared with ~1 kcal mol^{-1} in methanolic solution. Accordingly, the barriers

found in water might well account for kinetic control of complex formation, even if tiny interaction energies are found.

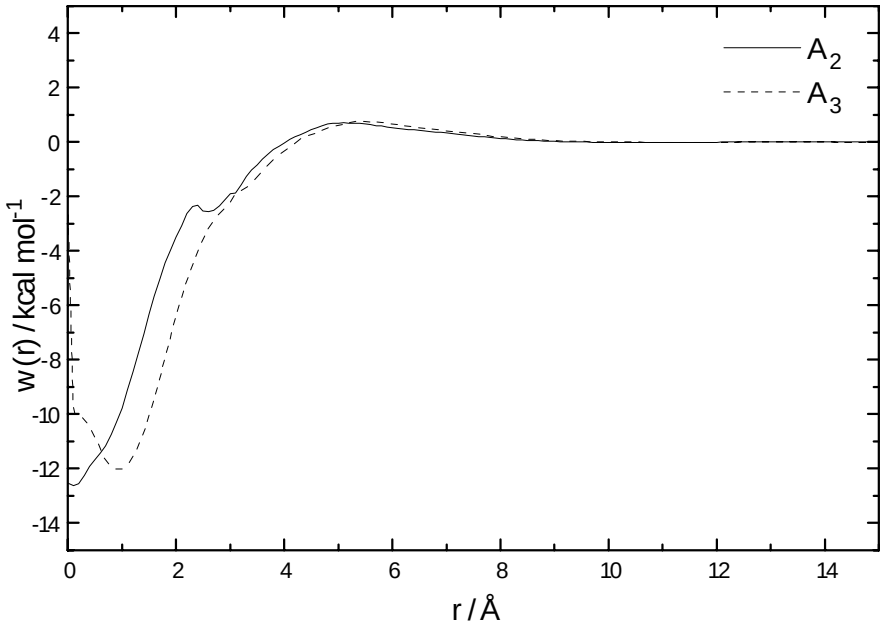

Figure 11: Potentials of mean force for the anchoring of amines A_2 and A_3 to 18-crown-6 in methanol. The minima of the PMFs are deep and broad in comparison to water, signalizing a weaker interaction of the guest with the solvent.

TABLE 2: Thermodynamic data and standard-state corrections SSC for the interaction of organic ammonium hosts with 18-crown-6 observed in aqueous and methanolic solution. The correction for the standard state is calculated according to eq 3.16.

Solvent	**Water**			**Methanol**	
Guest	**A_1**	**A_2**	**A_3**	**A_2**	**A_3**
Dissociation limit / Å	0.6	0.6	1.3	3.5	4.0
Exp. $\Delta G^{o\,2,3}$ / kcal mol^{-1}	-1.4 ± 0.2	-0.8 ± 0.4	-0.2	-5.45	-5.86
Calc. ΔG / kcal mol^{-1}	-0.16	-0.6	0.0	-5.05	-5.74
SSC / kcal mol^{-1}	-3.27	-2.44	-2.44	-2.44	-2.44
Calc. ΔG^{o} / kcal mol^{-1}	-3.43	-3.04	-2.44	-7.49	-8.18

Eq 3.15 was used to calculate free energies of binding from the potentials of mean force that were corrected for the standard state 1 mol l^{-1} by use of eq 3.16. The key thermodynamic results from the simulations are compiled in Table 2, wherein the dissociation limits of the complexes were estimated from the region where the potentials of mean force become positive. It is found, that the relative order of the calculated free energies of complexation matches the experimental results, while the calculated standard free energies show a constant offset of approximately 2.5 kcal mol^{-1} compared with experimental results.

3.5.3 Complexation of crown ethers with urea and guanidinium derivatives

Recently, thermodynamic data for the interaction of urea (A_4) and some guanidinium derivatives (A_4, A_5, A_6, cf. Fig 1) with crown ethers dissolved in methanol have been compiled. Buschmann and coworkers[10] reported log*K* data that are of the same order of magnitude as the binding constants determined with R-NH_3^+/18-crown-6. Clearly, the binding modes of -NH_2 residues provided by urea, guanidinium and its derivatives $R_4N_2CNH_2^+$ differ substantially from the anchoring mode of R-NH_3^+. However, the number of hydrogen bonds formed between host and guest cannot be obtained from the measured reaction enthalpies.[10]

The complexation reactions are mainly driven by electrostatic interactions between the ether oxygen atoms and the polar hydrogens of the amide groups. We used the methods described in the previous section to obtain minimum energy conformations of the various complexes in solution. An overview of the calculated structures is given in fig 12. Obviously, more than one ether oxygen of 18-crown-6 is involved in hydrogen bonding to the guanidinium-type-molecules. Urea, guanidine and benzamidine adopt a minimum energy conformation which allows electrostatic interaction even between the second -NH_2 group and an ether oxygen. In their complexes with 15-crown-5, however, the -NH_2 groups act as an anchor binding simultaneously to neighbored ether oxygens of the host. On the other hand, A_6 binds similarly to 18-crown-6 and 15-crown-5 by a single hydrogen bond, while the hydrophobic interaction with the solvent is minimized by screening of one of the methyl groups.

Thermodynamics of the crown-ether-guest complexation have been calculated with our RISM/MC algorithm. The experimental and calculated free energies are given in Tables 3 and 4. In accordance to experimental findings, the free

energies of complexation are all of the same order of magnitude. Also the *relative order* of the data mostly matches the experimental data. However, we note a general tendency of the method to overestimate standard free energies in comparison to experimental data. A proper sampling of all intramolecular degrees of freedom during the MC docking process would probably account for the differences observed, but multiply the calculation times spent in sampling of the solute-solute statistics.

TABLE 3: **Thermodynamic parameters calculated for the interaction of guanidinium-type guest molecules with 15-crown-5. Experimental findings as reported by Buschmann and coworkers[10] are given.**

15-crown-5	**A_4**	**A_5**	**A_6**	**A_7**
Dissociation limit / Å	5.8	5.0	5.3	5.5
Exp. ΔG^o / kcal mol^{-1}	-3.34	-5.01	-3.46	-4.39
Calc. ΔG / kcal mol^{-1}	-4.68	-4.64	-4.07	-4.97
SSC / kcal mol^{-1}	-2.19	-2.44	-2.19	-1.79
Calc. ΔG^o / kcal mol^{-1}	-6.87	-7.08	-6.26	-6.76

TABLE 4: Compilation of all thermodynamic parameters and standard-state corrections SSC for the interaction of guanidinium-type guest molecules with 18-crown-6 in comparison to experimental data.[10]

18-crown-6	**A_4**	**A_5**	**A_6**	**A_7**
Dissociation limit / Å	4.9	4.9	5.3	5.9
Exp. ΔG^o / kcal mol^{-1}	-3.34	-5.80	<-6.81	5.16
Calc. ΔG / kcal mol^{-1}	-4.65	-4.90	-4.55	-4.87
SSC / kcal mol^{-1}	-2.84	-3.08	-2.84	-2.44
Calc. ΔG^o / kcal mol^{-1}	-7.49	-7.98	-7.39	-7.31

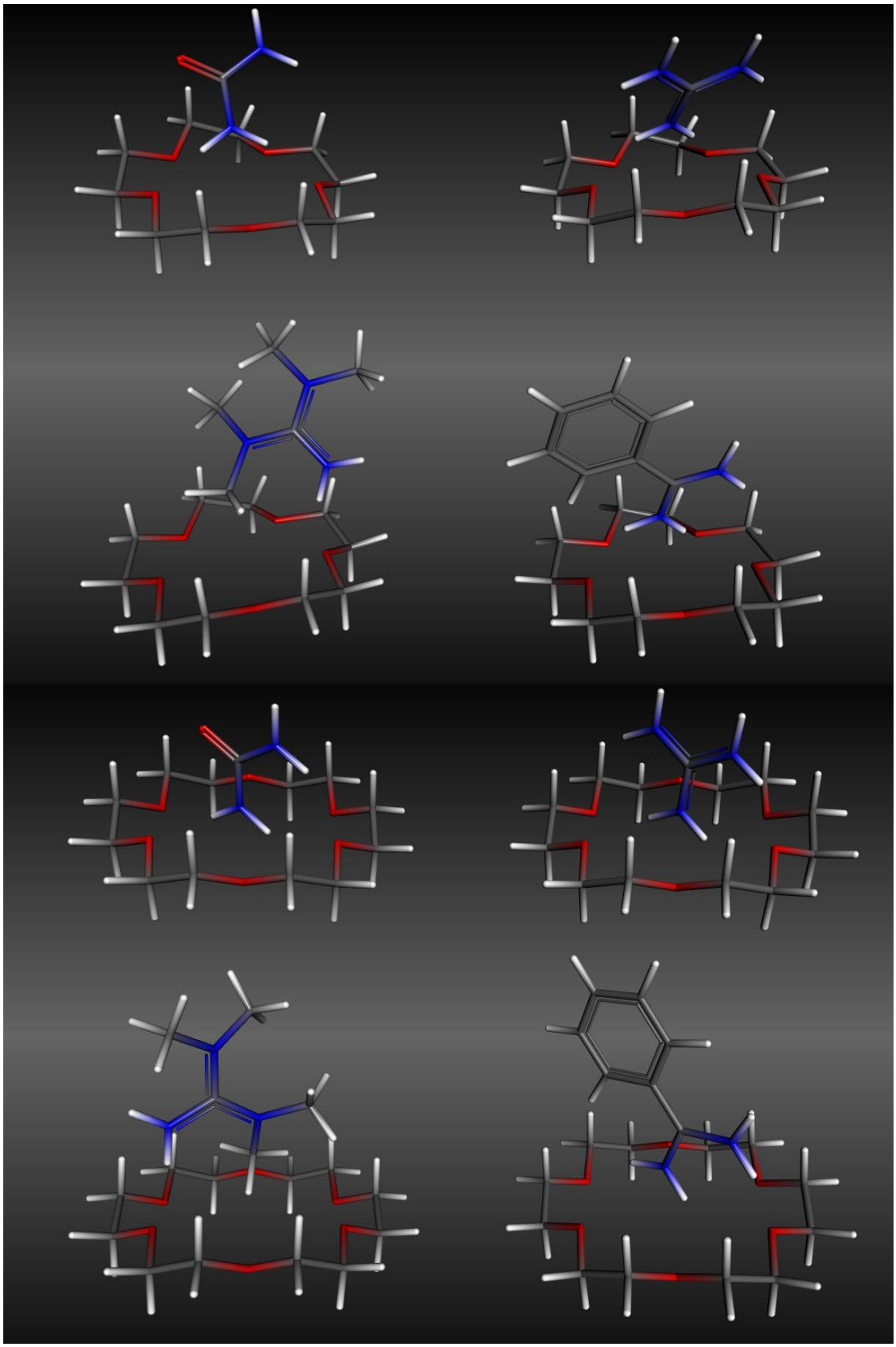

Figure 12: Global minimum energy conformers of the $A_{4...7}$/15-crown-5 and $A_{4...7}$/18-crown-6 complexes in methanol.

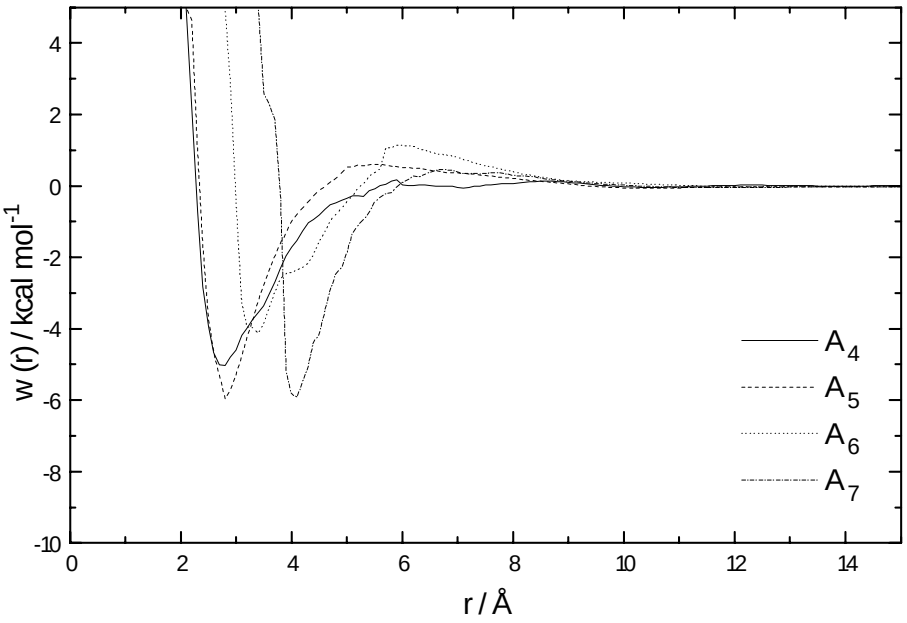

Figure 13: Potentials of mean force for the binding of guanidinium derivatives A_4 to A_7 to 15-crown-5 in methanolic solution. The reaction coordinate corresponds to the separation of the centers-of-mass. Little barriers are observed, that separate a clearly defined binding position from the nonbonded state.

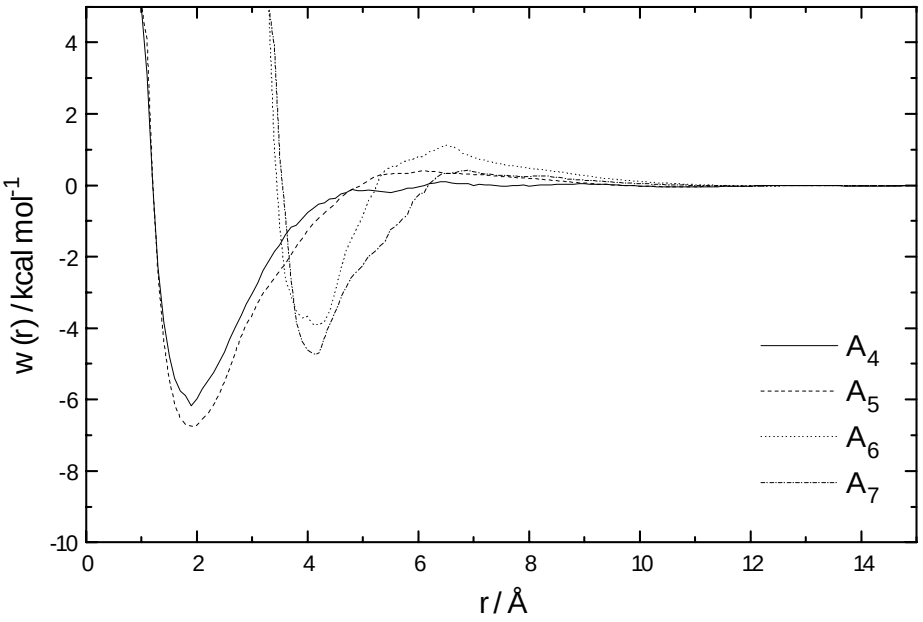

Figure 14: Potentials of mean force for the binding of guanidinium derivatives A_4 to A_7 to 18-crown-6 in methanolic solution. Again, the reaction coordinate corresponds to the separation of the centers-of-mass. The binding energies are comparably to 15-crown-5.

3.6 Conclusions

We have reported results of Monte Carlo simulations of solute-solute statistics concerning the complexation of various organic ammonium cations with 18-crown-6 and 15-crown-5 in solution. The solvents water and methanol were modeled by means of a fast statistical integral equation representation. Our RISM/MC interaction site approach, which was primarily designed to describe the interaction of metal ions with a crown-ether host in a polar solvent, was extended to be applicable to the complexation of arbitrary molecules. It was found that the integral equation approach is capable to provide structural features of the solvation structure around the host 18-crown-6.

RISM integral equation theory has been combined with a simplex minimization of a fully flexible all-atom force field representation of molecular complexes. Due to the inclusion of the solvation free energy in the total energetics it allows the fast determination of molecular structures in solution under more realistic conditions than generally provided by vacuum force fields. A detailed study of the interaction free energy surfaces of $H_3CNH_3^+$/18-crown-6 in water and methanol reveals symmetric minimum energy pathways which involve the step-by-step formation of three hydrogen bonds. These pathways are short-ranged due to the strong dielectric shielding of the solvents. Interference of the solvent shells arises long-ranged ordering of the supramolecular assembles. Specifically in water, the preferential orientation of the guest molecules dipole moment is parallel to the surface of the host and reminds to the hydration structure of a hydrophobic host.

Minimum energy structures of urea/18-crown-6 and some guanidinium derivatives in methanol reveal evidence for the concurrent interaction of two to three polar hydrogen atoms of the guest with the ring oxygens of 18-crown-6. In complexes with the smaller 15-crown-5, the same guests form two hydrogen bonds that remind to the anchoring mechanism discussed before.

By use of the hybrid RISM/MC algorithm presented in the first section, we have calculated free energies of complexation for the various complex partners in the infinite dilution limit. The results were corrected for standard state condition and were shown to be in good agreement with the experimental data.

Ongoing work is concerned with the topic of discrimination of chiral ammonium guests by crown ether compounds in solution, where more elaborate techniques will be needed to handle the increasingly complex FES of larger systems.

3.7 References

(1) Roberts, J.L.; McClintock, R.E.; El-Omrani, Y.; Larson, J.W. *J. Chem. Eng.* **1979**, *2*, 79.

(2) Izatt, R.M.; Bradshaw, J.S.; Nielsen, S.A.; Lamb, J.D.; Christensen, J.J.; Sen, D. *Chem. Rev.* **1985**, *85*, 271.

(3) Izatt, R.M.; Pawlak, K.; Bradshaw, J.S.; Bruening, R.L. *Chem. Rev.* **1991**, *91*, 1721.

(4) Izatt, R.M.; Bradshaw, J.S.; Pawlak, K.; Bruening, R.L. *Chem. Rev.* **1992**, *92*, 1261.

(5) Izatt, R.M.; Pawlak, K.; Bradshaw, J.S.; Bruening, R.L. *Chem. Rev.* **1995**, *95*, 2529.

(6) Buschmann, H.-J.; Schollmeyer, E.; Mutihac, L. *Supramol. Sci.* **1998**, *5*, 139.

(7) Hasani, M.; Shamsipur, M. *J. Solution Chem.* **1994**, *23*, 721.

(8) Gehin, D.; Kollman, P.A.; Wipff, G. *J. Am. Chem. Soc.* **1989**, *111*, 3011.

(9) Ha, Y.L.; Chakraborty, A.K. *J. Phys. Chem.* **1993**, *97*, 11291.

(10) Buschmann, H.-J-; Dong, H.; Mutihac, L.; Schollmeyer, E. *J. Therm. Anal. Cal.* **1999**, *57*, 487.

(11) Glendening, E.D.; Feller, D.; Thompson, M.A. *J. Am. Chem. Soc.* **1994**, *116*, 10657.

(12) Feller, D.; Apra, E.; Nichols, J.A.; Bernhold, D.E. *J. Chem. Phys.* **1996**, *105*, 1940.

(13) Dang, L.X.; Kollman; P.A. *J. Am. Chem. Soc.* **1990**, *112*, 5716.

(14) Auffinger, P.; Wipff, G. *J. Am. Chem. Soc.* **1991**, *113*, 5976.

(15) Dang, L.X.; Kollman, P.A. *J. Phys. Chem.* **1995**, *99*, 55.

(16) Varnek, A.; Wipff, G. *J. Comp. Chem.* **1996**, *17*, 1520.

(17) Jorgensen, W.L.; Bucker, J.K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742.

(18) Pranata, J.; Jorgensen, W.L. *Tetrahedron* **1991**, *47*, 2491.

(19) Nguyen, T.B.; Jorgensen, W.L. *Proc. Natl. Acad. Sci.* **1993**, *90*, 1194.

(20) Du, Q.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **2000**, *104*, 796.

(21) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10391.

(22) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10403.

(23) Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **1999**, *103*, 7942.

(24) Exner, T.E.; J. Brickmann, J. *J. Mol. Model.* **1997**, *3*, 321.

(25) Jäger, R.; Schmidt, F.; Schilling, B.; Brickmann, J. *J. Comp.-Aided Mol. Des.* **2000**, *14*, 631.

(26) Schmidt, K.F.; Kast, S.M. *J. Phys. Chem. B*, submitted for publication.

(27) Hansen, J.P..; McDonald, I.R. *Theory of simple liquids*, Academic Press, London **1980**.

(28) Melenkevitz, J.; Schweitzer, K.S.; Curro, J.G. *Macromolecules* **1993**, *26*, 6190.

(29) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Am. Chem. Soc.* **1998**, *120*, 1855.

(30) Khalatur, P.G.; Khokhlov, A.R. *Mol. Phys.* **1998**, *93*, 555.

(31) Prue, J.E. *J. Chem. Educ.* **1969**, *46*, 12.

(32) Hirata, F.; Rossky, P.J. *Chem. Phys. Lett.* **1981**, *83*, 329.

(33) Chandler, D.; Andersen, H.C. *J. Chem. Phys.* **1972**, *57*, 1930.

(34) Singer, S.J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

(35) Karplus, M. CHARMM Documentation, The CHARMM Development Project, Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

(36) SYBYL 6.6, *Tripos Inc.*, 1699 South Hanley Road, St. Louis, Missouri, 63144 USA.

(37) Clark, M.; Cramer III, R.D.; van Opdenbosch, N. *J. Comp. Chem.* **1989**, *10*, 982.

(38) Numerical Recipes in C: THE ART OF SCIENTIFIC COMPUTING, Cambridge University Press, **1988**, 408.

(39) Troxler, L.; Wipff, G. *J. Am. Chem. Soc.* **1994**, *116*, 1468.

(40) Kowall, T.; Geiger, A. *J. Phys. Chem.* **1995**, *99*, 5240.

(41) Gilson, M.K.; Given, J.A.; Bush, B.L.; McCammon, J.A. *Biophys. J.* **1997**, *72*, 1047.

(42) Hill, T.L. *Statistical Thermodynamics*, Addison-Wesley Publishing, London **1950**.

(43) Jorgensen, W.L. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(44) Pettitt, M.B.; Rossky, P.J. *J. Chem. Phys.* **1986**, *84*, 5836.

(45) Jorgensen, W.L. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(46) The Cambridge Structural Database, Cambridge Crystallographic Data Center, 12 Union Road. Cambridge. CB2 1EZ. UK.

(47) Talman, J.D. *J. Comput. Phys.* **1978**, *29*, 35.

(48) Rossky, P.J.; Friedman, H.L. *J. Chem. Phys.* **1980**, *72*, 5694.

(49) Patil, K.; Pawar, R. *J. Phys. Chem. B* **1999**, *103*, 2256.

(50) Takeda, Y.; Kanazawa, M.; Katsuta, S.; *Anal. Sci.* **2000**, *16*, 929.

(51) Rhangino, G.; Romano, S.; Lehn, J.M.; Wipff, G. *J. Am. Chem. Soc.* **1985**, *107*, 7873.

(52) Straatsma, T.P.; McCammon, J.A. *J. Chem. Phys.* **1989**, *91*, 3631.

(53) Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C.D.; *Interactive Visualization of Molecular Scenarios with the MOLCAD/SYBYL Package*. In: Bowie, J.E.; Ed.; Data Visualization in Molecular Science "Tools for Insight and Innovation"; Addison-Wesley Publishing Company Inc., Reading, Mass. **1995**.

(54) Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G.*; Molecular Models: Visualization.* In: Schleyer, P.v.R.; Allinger, N.L.; Clark, T.; Gasteiger, J.; Kollman, P.A.; Schaefer III, H.F.; Schreiner, P.R., Eds.; "The Encyclopedia of Computational Chemistry", John Wiley & Sons: Chichester, **1998**.

4 Theoretical calculation of chiral selection in ion-crown ether complexes with a hybrid integral equation theory/genetic algorithm approach

4.1 Summary

Computational approaches to complexation thermodynamics in solution commonly employ detailed microscopic solvent models in expensive free energy simulations. Most of the computational effort is spent in the sampling of the solvent molecules. For rapid estimation of thermodynamic observables other techniques such as the reference interaction site model RISM are preferable. In principle, this theory generally provides a qualitatively correct description of the site-site distribution functions in a molecular liquid. The structural information is included in terms of site-site distance matrices. Consequently, chiral molecules are found to be identical with their mirror images, making it difficult to observe thermodynamic interaction data of chiral molecules within the RISM method. To this end, we have combined the RISM method for the solvent phase with sampling techniques for the solute degrees of freedom to explore the free energy surface (FES).

A genetic algorithm in combination with integral equation theory is employed to explore the topography of the FES with respect to locating the global minimum. Therein, constrained local optimization is used to enhance the performance of the genetic algorithm. The relative stability of enantiomeric host-guest complexes, commonly known as chiral selection, is expressed as the difference of the global minima of the FES, $\Delta\Delta G_{R\text{-}S}$, between enantiomeric complexes and is compared with experimental data. The hybrid approach is employed to calculate chiral selections of 10 organic ammonium-ion/pyridino-18-crown-6 complexes in methanolic solution.

4.2 Introduction

Computational studies of molecular recognition phenomena[1] have received considerable attention over the last decades. Much effort has been spent in the

development of liquid state simulation techniques[2,3,4] for the simulation of biologically active molecules.[5,6] Much of the research is also directed in the design of new pharmacophores or agrochemicals, where molecular chirality is often an important factor in determining the activity of a specific compound. Whenever asymmetric synthesis fails, enantioseparation is of wide importance for the purification of compounds in technical chemistry. Several classes of chiral selectors have been used for that purpose, including proteins[7], cyclodextrines[8] and macrocycles[9]. Chromatographic methods[10] and capillary zone electrophoresis[11] take advantage of the high enantiodiscriminating powers of chiral macrocyclic compounds[12]. Crown ethers have been immobilized to form HPLC phases which form complexes with protonated primary amines. A commonly used crown ether has been 18-crown-6 which was also made available commercially[13]. Crown ethers often give efficient separations allowing the reversal of elution order in cases of trace analysis of the unwanted enantiomer.

Particularly, the recognition of chiral ammonium ions by pyridine-containing crown ethers is known for strong discriminating forces[14]. This system is well suited for docking studies, since the complex structures and thermodynamics have been measured exhaustively[12].

The recognition process of a chiral compound involves stereochemical interactions with an enantioselector, both forming a diastereomeric complex with six centers interacting simultaneously[15]. The thermodynamic stability of a given complex is classified by its average free energy of complexation, ΔG. The relative stability of a set of complexes (e.g. a diastereomeric pair of complexes) might be identified with the difference of their free energies of complexation, $\Delta\Delta G_{R\text{-}S}$[16]. The difference in the complexation energies of a chiral host with the enantiomeric forms of a guest molecule is termed chiral selection.

Our work has focused on the development of a hybrid algorithm which is capable to calculate differential thermodynamics of host-guest interactions in solution[17]. In a preliminary approach, RISM integral equation theory has been combined with Monte Carlo simulation of solute-solute statistics concerning the complexation of metal ions[18,] and organic ammonium cations[19] with 18-crown-6 and 15-crown-5 in solution. The RISM/MC method was successfully used to calculate free energies of complexation. However, the more complex free energy surface of

chiral host-guest complexes requires the development of more elaborate docking techniques.

In this study, the reference interaction site model (RISM) integral equation theory[20,21] is applied to derive the solution thermodynamics. This integral equation, though not exact, has been shown to yield a reasonable picture of molecular fluids[22] while providing free energy expressions[23,24] which can be implemented in atomistic simulation methods.[18,19]

Within the RISM model, the molecular structural is treated in terms of site-site distance matrices. This description cancels some information concerning the chirality of a molecule. The same distance matrix is used to describe chiral molecules and their mirror images, making it difficult to observe thermodynamic interaction data of chiral molecules within the RISM method[25,26]. Therefore, an explicit molecular treatment of solute-solute interaction is useful. To this end, RISM is combined with a genetic algorithm (GA) with respect to location the global minimum of chiral host-guest-complexes. In contrast to preliminary studies with the RISM/MC technique,[18,19] the present method focuses on molecular docking in solution by means of a free-energy-scoring function.

This paper is organized as follows. In the following section we present the methods and algorithmic details related to the hybrid RISM/LGA method. Section 3 provides computational details. In chapter 4, RISM/LGA is applied to calculate chiral selection data of organic ammonium-ion/crown ether complexes in methanolic solution. Concluding remarks are given in the final section.

4.3 Methods

This section gives descriptions of the computational procedures that are used in constructing the hybrid RISM/LGA algorithm. The details of the stand-alone genetic algorithm, RISM integral equation implementations, and a local search algorithm are outlined below. This is followed by a subsection exploring the merger of the algorithms and its application to molecular docking.

4.3.1 Genetic Algorithm

The idea of applying fundamental biological principals of natural evolution to global optimization problems in science was pioneered by Holland[27] in 1975.

Genetic algorithms (GA)[28] use computational models of natural evolutionary processes as tools in the design and implementation of global optimization problems[29]. The concept is based on simulating the evolution of a population via processes of reproduction and selection, which are applied to the individuals forming the population (fig. 1). The fitness of an individual is used as a quantitative measure to control the perceived performance of the individual during selection and reproduction.

After random initialization of the population, the number of individuals is kept constant during the evolution process. Offspring is created via reproduction, which conserves the individual's information, recombination, which exchanges information between the parents, and mutation, which alters the information of individual children.

The Darwinian selection process involves the use of a fitness function, focusing attraction to the (n-2) fittest members of the population, which are selected for breeding. The fittest individual of the population is copied to the next generation by use of the reproduction operator. The offspring of the breeding parents is created via recombination of the genetic information, coded binarily in strings containing all necessary structural parameters (genome). In the recombination process, two parents are probabilistically selected from the (n-2) fittest members of the population. A crossover point is randomly chosen in the genomes of the parents. The children are generated subsequently by an interchange of the genome subsets defined by the crossover section between the parents.

In the mutation operation, a single individual from the offspring is probabilistically chosen based on the individual's fitness. A mutation spot is chosen, and the residual genomic information is randomly reconstructed according to the rules used in the random initialization of the initial population.

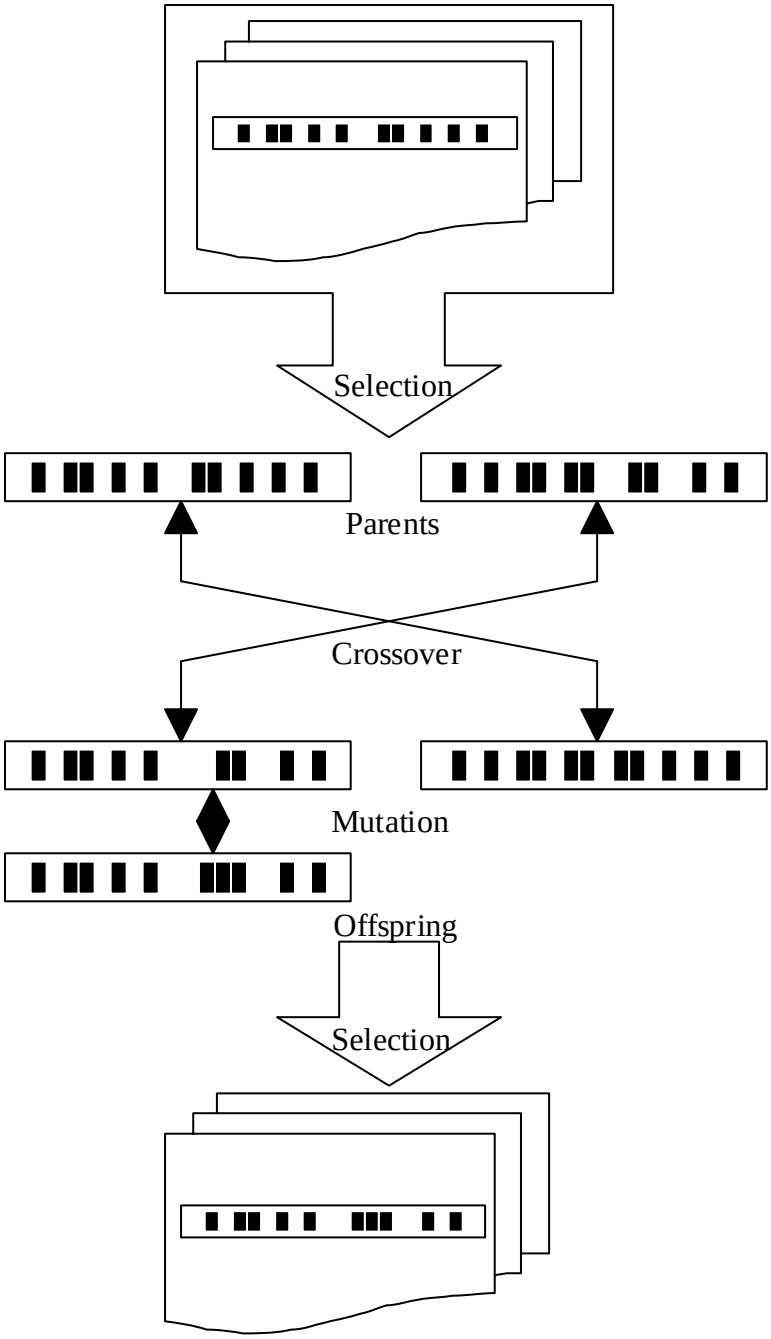

Figure 1: Flow chart of the genetic algorithm with binary coding of the chromosomes. The genetic operators are crossover, mutation and reproduction.

4.3.2 Local optimization

Combinations of genetic algorithms with local search (LS) strategies have been investigated by many researchers[30,31,32,33]. It was shown that the effectiveness of GA algorithms for solving combinatorial optimization problems is enhanced significantly performing local search for each individual in the entire population[30]. In

fact, the implementation of local search allows the genetic algorithm to restrict the search to the space of local minima on the FES rather than performing a search over the entire FES.

The reliability of the fitness function, which is used to score conformers produced by the genetic algorithm, is essential to the optimization results. Our fitness function bases on the all-atom force field representation of TRIPOS inc.[34,35] accounting for intramolecular interactions. A configurational dependent solvent contribution $\Delta\mu^{solv}$ was added based on integral equation theory. The force field energy of a molecular ensemble in vacuum E^{vac} reads

$$\begin{aligned} E^{\text{vac}} = & \sum_{\text{i}}^{\text{bonds}} \frac{1}{2} k_{\text{i}}^{\text{bond}} \left(d_{\text{i}} - d_{\text{i}}^{\text{o}}\right)^2 + \sum_{\text{j}}^{\text{angles}} \frac{1}{2} k_{\text{j}}^{\text{angle}} \left(\theta_{\text{j}} - \theta_{\text{j}}^{\text{o}}\right)^2 \\ & + \sum_{\text{l}}^{\text{torsions}} \frac{1}{2} V_{\text{l}}^{\text{torsion}} \left[1 + S_{\text{l}} \cos\left(\left|n_{\text{l}}\right| \cdot \vartheta_{\text{l}}\right)\right] \\ & + \sum_{\text{m}}^{\text{out-of-plane}} \frac{1}{2} k_{\text{m}}^{\text{oop}} d_{\text{m}}{}^2 + E^{\text{nonbonded}} \end{aligned} \tag{4.1}$$

where k_{i}, k_{j}, k_{m} represent the bond stretching, angle bending and out-of-plane bending constants, respectively. Equilibrium values for the bondlength and the angles are termed d_{i}^{o} and $\theta_{\text{j}}^{\text{o}}$. V_{l} is the torsional barrier of the l^{th} torsion, n_{l} is its periodicity and ϑ_{l} is the torsion angle. S_{l} equals to one, if the minimum energy of the torsion is found in staggered conformation, and equals to -1, if the eclipsed conformation is preferred. Hydrogen bonding terms are included implicitly in the force field parameters.

The force field parameters do not account for solvent contributions to the total energy of the system. Therefore, a conformational-dependent solvent contribution $\Delta\mu^{solv}$ was included from statistical mechanical integral equation theory[23]. The total energy E^{tot} of a molecular ensemble in a RISM solvent reads

$$E^{\text{tot}} = E^{\text{vac}} + \Delta\mu^{\text{solv}} \ . \tag{4.2}$$

By this means, the orientationally dependent potential of mean force $W(\mathbf{\Gamma})$ between host and guest molecules[18] is given by eq 4.3:

$$\begin{aligned} W(\mathbf{\Gamma}_{\text{XY}}) = & \Delta V_{\text{XY}}(\mathbf{\Gamma}_{\text{XY}}) + E_{\text{XY}}^{\text{vac}}(\mathbf{\Gamma}_{\text{XY}}) + \Delta\mu_{\text{XY}}(\mathbf{\Gamma}_{\text{XY}}) \\ & - \left[E_{\text{X}}^{\text{vac}}(\mathbf{\Gamma}_{\text{X}}) + \Delta\mu_{\text{X}}(\mathbf{\Gamma}_{\text{X}})\right] - \left[E_{\text{Y}}^{\text{vac}}(\mathbf{\Gamma}_{\text{Y}}) + \Delta\mu_{\text{Y}}(\mathbf{\Gamma}_{\text{Y}})\right] \end{aligned} \tag{4.3}$$

$\mathbf{\Gamma}_X$, $\mathbf{\Gamma}_Y$, and $\mathbf{\Gamma}_{XY}$ are the coordinates of the host X, the guest Y and the complex XY, respectively. The vacuum interaction energy is denoted by ΔV_{XY}. If applied to the calculation of rigid body interaction only, eq 4.3 simplifies considerably:

$$W^{\text{rigid}}(\mathbf{\Gamma}_{XY}) = \Delta V_{XY}(\mathbf{\Gamma}_{XY}) + \Delta\mu_{XY}(\mathbf{\Gamma}_{XY}) - \Delta\mu_X(\mathbf{\Gamma}_X) - \Delta\mu_Y(\mathbf{\Gamma}_Y) \qquad (4.4)$$

The fitness functions eqs 4.3 and 4.4 are of central importance in the optimization procedures.

Eqs 4.3 and 4.4 open the possibility to perform local optimization procedures of molecular ensembles in solution. A robust simplex downhill algorithm[36] is employed to perform local optimization procedures of complexes in a RISM solvent. A simplex is the geometrical figure in *n* dimensions consisting of *n*+1 vertices. The simplex algorithm for minimization takes such a set of *n*+1 points and uses reflection, expansion and contraction steps to move them into a local minimum. (fig. 2).

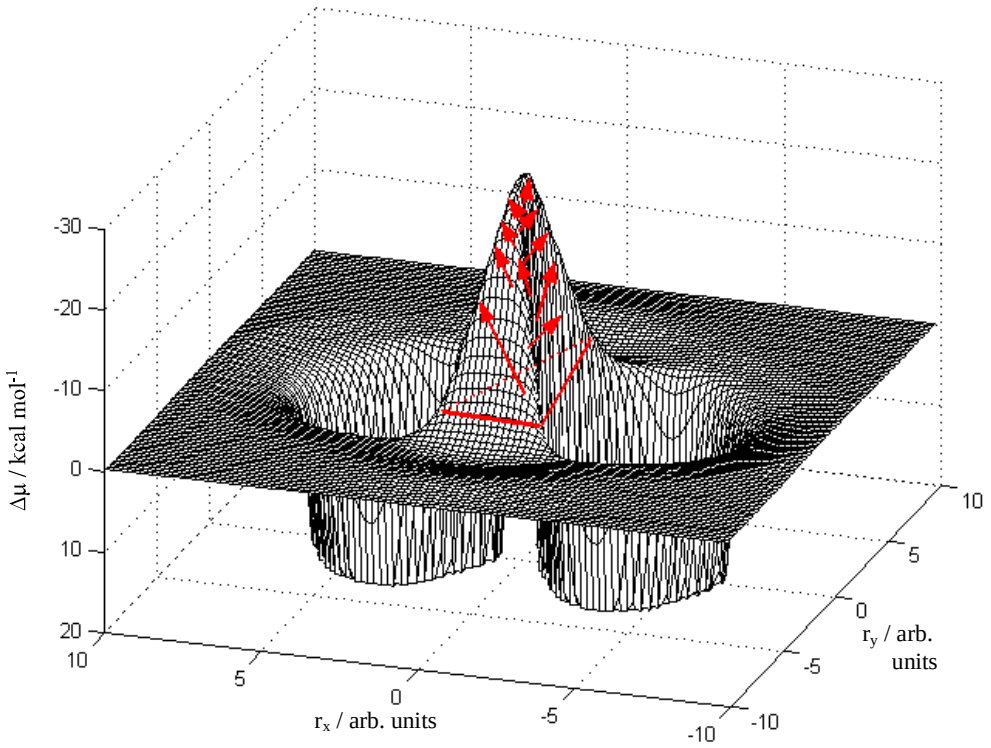

Figure 2: Schematic depiction of a 2-dimensional function minimization with a Simplex algorithm. The x and y axes are given in arbitrary units while is given in units of kcal mol^{-1}. The z axis has been swapped for clarification.

Taking the initial guess as $\mathbf{r}_o$, the simplex is initialized as

$$\mathbf{r}_n = \mathbf{r}_o + \lambda_n \mathbf{e}_n \tag{4.5}$$

where the **e** are unit vectors and λ is 0.1 (Å) for cartesian optimizations and 5 (degrees) for angular coordinates. The calculation terminates, when either the maximum number of steps is exceeded or the minimum has been located.

4.3.3 RISM integral equations

The details of the RISM integral equation theory were given in a preceding publication.[18] Within the RISM model, molecules are assumed to consist essentially of atomic interaction sites which are rigidly bonded. Each interaction site is associated with a set of intermolecular interaction parameters, which are taken from standard force fields.[34] The pairwise site-site interaction potential $U_{ij}(r)$, which is the sum of a short-ranged Lennard-Jones term and a long-ranged Coulomb term, reads:

$$U_{ij}(r) = U_{ij}^{\mathrm{LJ}}(r) + U_{ij}^{\mathrm{C}}(r) = 4\pi\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \kappa\frac{q_i q_j}{r_{ij}} \tag{4.6}$$

The partial charges of the molecular interaction sites in atomic units are denoted by *q*. The Lorentz-Berthelot combination rules are used to obtain interatomic short-range parameters σ_{ij} and ε_{ij}. The intermolecular Coulombic potential is corrected to yield an accurate screening of the long range interaction[37]:

$$\kappa = \begin{cases} \left[\dfrac{(\varepsilon_r - 1)kT + \frac{4}{3}\varepsilon_o \pi\rho\, p^2}{(\varepsilon_r - 1)\cdot\frac{4}{3}\pi\rho\, p^2}\right] & ;\ \text{solvent} - \text{solvent interaction} \\ 1 & ;\ \text{solute} - \text{solvent interaction} \end{cases} \tag{4.7}$$

where *T* is the absolute temperature, *p* is the absolute value of the dipole moment of the solvent molecule, *k* is the Boltzmann constant, ρ is the number density of the solvent methanol, ε_r the phenomenological dielectric constant of the solvent (ε_r = 32.66) and ε_o is the vacuum dielectric constant.

The RISM integral equation for the solute-solvent correlation functions $\mathbf{h}^{\mathrm{Xv}}$ reads

$$\mathbf{h}^{\mathrm{Xv}} = \boldsymbol{\omega}^{\mathrm{X}} * \mathbf{c}^{\mathrm{Xv}} * \boldsymbol{\chi}^{\mathrm{vv}} \tag{4.8}$$

where the matrix of intramolecular pair distribution functions $\boldsymbol{\omega}^{X}$ and the correlation function matrix $\mathbf{h}^{Xv}$ are X · X and X · v labeled matrices, respectively. The star (*) denotes matrix convolution products. Accounting for a given molecular conformation, $\boldsymbol{\omega}^{X}$ consists of delta functions, which constrain the atomic interaction sites *i*, *j* to the bondlength $|\mathbf{d}_{ij}|$:

$$\omega_{ij}(\mathbf{r}) = (1-\delta_{ij})\frac{\delta(|\mathbf{r}|-|\mathbf{d}_{ij}|)}{4\pi|\mathbf{d}_{ij}|^2} \tag{4.9}$$

The solvent-solvent susceptibility matrix $\boldsymbol{\chi}^{vv}$

$$\boldsymbol{\chi}^{vv} = \boldsymbol{\omega}^{v} + \rho\,\mathbf{h}^{vv} \tag{4.10}$$

depends on the number density of the solvent and is invariant under insertion of an arbitrary solute molecule. It is derived by a pure solvent calculation and can be taken directly from an input file.

The hypernetted chain closure (HNC)[38] for the matrix elements *ij* is written as

$$h_{ij} = \exp\left(-\frac{U_{ij}}{kT} + h_{ij} - c_{ij}\right) - 1 \tag{4.11}$$

A complete set of distribution function as obtained by the solution of eqs 4.8 and 4.11 is needed to calculate the free energy of solvation[23] $\Delta\mu$ for a solute in a given conformation $\boldsymbol{\omega}^{X}$:

$$\Delta\mu(\boldsymbol{\omega}^{X}) = -\frac{\rho}{\beta}\sum_{i,j}\int d\mathbf{r}\left(\frac{1}{2}h_{ij}^2 - c_{ij} + \frac{1}{2}h_{ij}c_{ij}\right) \tag{4.12}$$

The difference in the free energy of solvation related to any structural change[24] X→X' is then given by

$$\Delta\Delta\mu(\boldsymbol{\omega}^{X},\boldsymbol{\omega}^{X'}) = -\frac{\rho\,kT}{2}\sum_{i,i'}\int d\mathbf{r}\left[\mathbf{c}^{Xv} * \boldsymbol{\chi}^{vv} * \mathbf{c}^{X'v}(\boldsymbol{\omega}^{X} - \boldsymbol{\omega}^{X'})\right]_{i,i'} \tag{4.13}$$

As has been shown in a previous publication[18], the use of the variational expression 13 in the determination of free energies of solvation greatly saves computational times.

4.3.4 GA-based docking with local search and a RISM solvent

The genetic algorithm with local search (LS/GA) is applied to the calculation of the complexation thermodynamics of chiral pyridine-containing crown-ether complexes in methanolic solution. Therein LS/GA is coupled to an RISM integral equation solver, which provides solvation thermodynamics derived from the radial distribution of the solvent around a complex (RISM/LGA).

In our automated docking procedure, the binding of a ligand to the host is modeled in a two-stage procedure. The first step is to place host and ligand molecules at a certain distance (e.g. 20 Å) apart and run the hybrid LS/GA algorithm for rigidly bonded complex partners. This pre-docking step corresponds to the scan of potential binding sites resulting in a set of local minima ready for flexible docking. In the second step, a fine-tuning of the complex structures is made, allowing for full relaxation of the internal coordinates. The separation of the docking process into rigid and flexible part has been recently used in the Glob-MolINE approach[39] and was found to improve the efficiency of the procedure.

```
begin
    solve RISM equations to obtain fitness function
    initialize individuals [I1,...IN] randomly;
    for j=1 to j=N Ij = execute function(local_optimization);
    do
       IN = execute function(elitism);
        for j=1 to N - 1
            {Iparent1;Iparent2} = execute function(select_parents_probabilistically);
            Ichild = execute function(crossover);
            if (random) Ichild = execute function(creep_mutation);
            if (random) Ichild = execute function(jump_mutation);
            Ichild = execute function(local_optimization);
        for j=1 to j=N-1 Ij = execute function(select_tournament);
    while (not converged)
end
```

Figure 3: Outline of procedures of the hybrid RISM/LGA algorithm for the pre-docking step.

The outline of the algorithmic procedures is given in fig 3. The hybrid algorithm works as follows. Initially, the RISM equations are solved for a reference complex in arbitrary geometry. The rigid body fitness function eq 4.4 is constructed. Now the entire population is initialized by random variation of intermolecular coordinates. The creation of an individual is followed by a local optimization

procedure, as has been described in the previous section. Doing this, a population of local minima within the RISM-based free energy surface is obtained. The elitism operator transfers the fittest individual immediately to the next generation unaltered. This is employed to ensure that the very best individual of each generation always survives and is not lost. The main loop of the algorithm starts with selection of N-1 individual pairs, which are chosen probabilistically according to the fitness criterion. To each pair, the crossover operator is applied. Additionally, mutation is applied randomly to a small number of the resulting individuals. Therein, creep mutation operations are controlled arithmetically to perform mutations which enhance the efficiency of the search by doing small changes in the gene value only. An equal number of jump mutations is applied which cause large changes in the chromosomes by random changing of bits in the chromosome string.

Subsequently, each child undergoes the local search optimization procedure. At the end of each loop, selection is performed to keep the number of local minima constant. Selection is now decided through tournament against another individuals. By this means, a fixed number of generations is crossbred. Once the calculation limits have been reached, the binary information is decoded into atomic coordinates, making the pre-docked complex structures subject to further investigation.

4.4 Model System

The achiral model system, pyridino-18-crown-6 **1**, is known to form strong hydrogen bonds with organic ammonium cations that are even stronger than those found in 18-crown-6 complexes[12]. Generally, the cation bonds preferentially to the nitrogen atom of the pyridine group of the host, while two other hydrogen bonds are formed to two alternate oxygen atoms of the crown ring. By introduction of side chains to the crown, stereocenters are formed and the host is enabled for the enantioselective recognition of chiral guest molecules. It has been found in solution phase experiments that in particular the recognition of α-(1-naphtyl)-ethylammonium by a dimethyldiketopyridino-18-crown-6 host is very strong[40]. Gas-phase enantiorecognition in this system has already been the subject of experimental[41] and computational[42] studies.

Chiral selections of organic ammonium ions by chiral pyridino-18-crown-6 derivatives were calculated in methanolic solution to study the performance of our docking method. Six host molecules including (S,S)-dimethylpyridino-18-crown-6 **2**,

(S,S)-dimethyldiketopyridino-18-crown-6 **3**, (S,S)-diphenyldiketopyridino-18-crown-6 **7** and derivatives of **2** and **3** were modeled using the molecular modeling package SYBYL[34] (fig 4).

All molecular guests (e.g. α-(1-naphtyl)ethylamine **G1**, phenylethylamine **G2**, methoxyalanine **G3**, cyclohexylamine **G4** and sec-butylamine **G5** were modeled in their protonated forms (fig 5). The stretch-, bend-, torsion-, improper torsion- and short-range interaction parameters were assigned according to the force field of TRIPOS inc.[34,35] All partial charges were taken from the OPLS all atom force field of Jorgensen and coworkers[43]. The structure and parameters of the solvent methanol were taken from the OPLS united-atom representation of rigid methanol.[44] For all polar hydrogen atoms, additional Lennard-Jones terms were added. σ_H was chosen as 0.4 Å - which is far less than σ_O or σ_N of the corresponding heavy atom. This modification has been proposed to avoid Coulomb singularity which otherwise arises inevitably during RISM calculations.[45] The additional potential has a negligible effect on the total interaction with the solvent.

All host and guest molecules were subjected to a relaxation of the intramolecular coordinates with the local optimization method. Accounting for solvent effects, the RISM equations 4.8 and 4.11 were solved for the starting structures on a logarithmic grid of 512 points, covering the interval [0.00598 Å; 164 Å]. For the calculation of Fourier transforms of the distribution functions the Talman method[46] was used. A damped iterative Picard procedure[38] was employed to achieve convergence of the distribution functions. The calculations were repeated until the maximum change in all direct correlation functions met the convergence criterion for the direct correlation function $\max\left(\Delta c_{ij}\right) < 10^{-5}$. Eq 4.2 was used to calculate free energies of solvation $\Delta\mu^{solv}$ for the starting structures.

The local search algorithm, which has been described in the preliminary section, was employed to calculate plausible lowest-energy structures of all hosts and guest molecules in solution. Therein, the total energy of the complex partners were minimized with our simplex downhill algorithm according to

$$E^{\mathrm{tot,min}}\left(\mathbf{\Gamma}_{\mathrm{X}}\right) = \min\left[E^{\mathrm{tot}}\left(\mathbf{\Gamma}_{\mathrm{X}}\right) + \Delta\mu^{\mathrm{solv}}\left(\mathbf{\Gamma}_{\mathrm{X}}\right)\right] \tag{4.14}$$

The minimum-energy structures in methanolic solution derived by this method were used as starting geometries for the rigid-body pre-docking step.

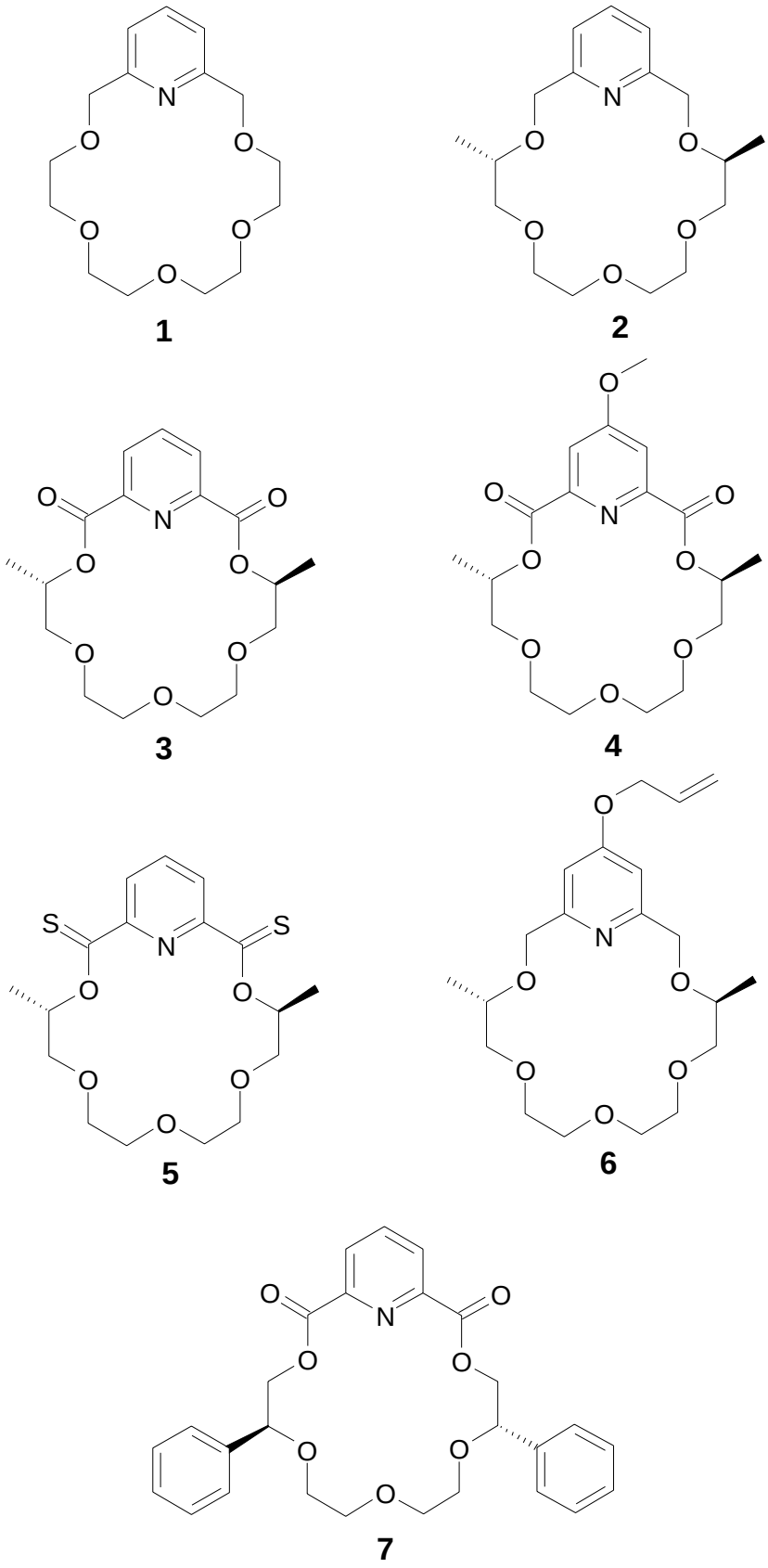

Figure 4: The model system pyridino-18-crown-6 (**1**) and the chiral macrocycles **2-7** examined in this study.

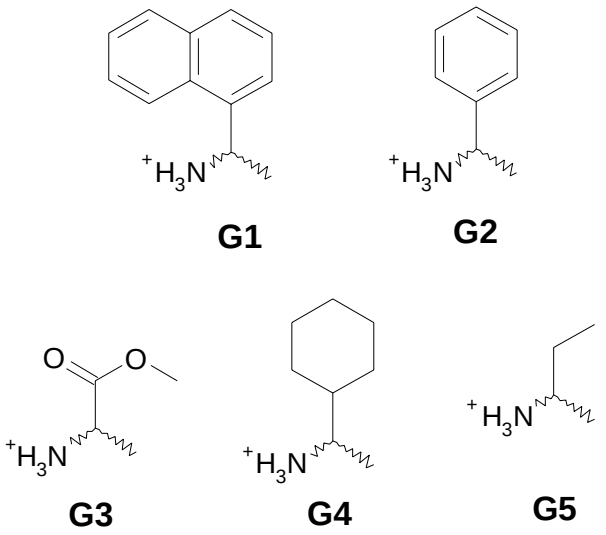

Figure 5: Organic ammonium guests examined in this study.

4.5 Results

4.5.1 Pre-Docking

The program GaFortran of Carroll[47] was used to perform the selection, crossover, mutation and tournament operations of binary strings. An interface to our integral equation solver was added in order to integrate the RISM based fitness function evaluating the total energy of the complexes. Local search was performed on the basis of the simplex downhill minimizer.

Initially, the complex partners were placed 20 Å apart into the RISM solvent. Again, the RISM equations were solved for the system to obtain a set of reference distribution functions. A reaction coordinate *r* was defined as the displacement of the centers-of-mass. The complexation energy surface of rigid bodies is spanned by six intermolecular coordinates[18], e.g. the reaction coordinate *r* and the rotation angles of the host $\{\theta, \varphi, \Theta\}$ and of the guest $\{\Phi, \Psi\}$. Both molecules were allowed to rotate freely, while the reaction coordinate span the range [0 Å;20 Å]. The parameters were given as floating point numbers. Due to binary coding, the minimum increment of each parameter depends on the length of the bit string used. In our implementation, the maximum number of possible adjustments of each variable was set to 2^{17}, giving a precision of 0.00275 degrees for the rotation angles and 0.00015 Å for the intermolecular distance. However, high precision does barely affect the calculation

times; but it is useful to distinguish between minima that are close in their structure but unequal in energy.

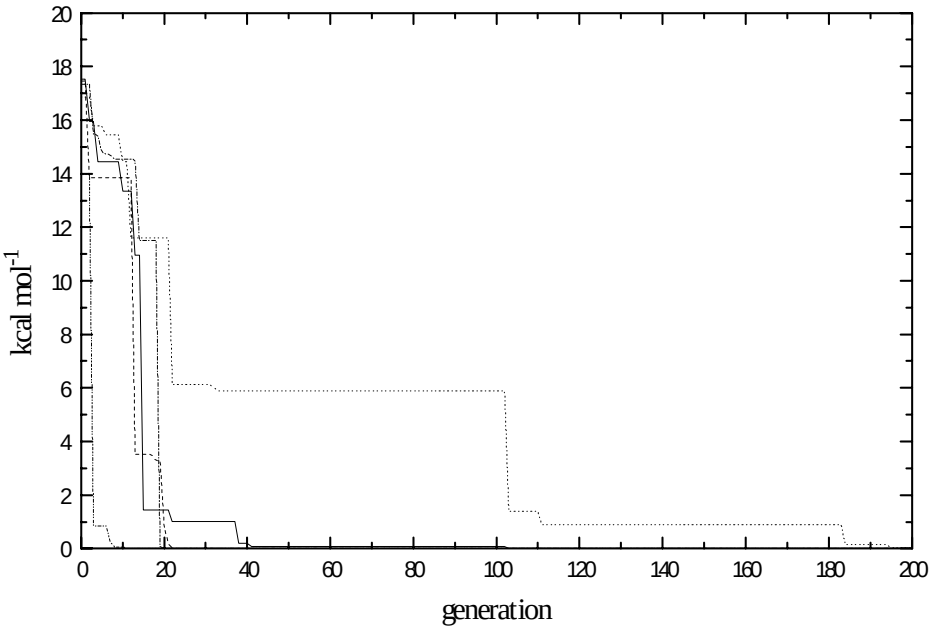

Figure 6: Runs of the RISM/LGA algorithm for the host (S,S)-dimethyldiketo-pyridino-18-crown-6 **3** with a protonated methoxyalanine guest **G3** in methanolic solution with different sets of initial conditions. The energy difference to the lowest-energy conformation is given. Most calculations are converged, after 50 generations have been evaluated.

A generation consisted of 10 individuals, that were subject to the genetic operators during the optimization process. The calculation parameters were chosen in accordance to the recommendations of ref.[47] .The crossover probability was set to 0.5 within the breeding process, e.g. two parents were selected and either recombined to form two children with a probability of 0.5 or allowed to pass unaltered to the next generation. The probabilities of creep and jump mutations were set to 0.04 and 0.05, respectively. The fitness of each individual complex was scored according to eq 4.3, and the complex was subjected to energetic minimization, resulting the total interaction energy. Therein, the binary information was decoded to structural information and a local optimization was made with an upper limit of 200 steps. The resulting structure was binarily recoded for tournament selection.

The evolution process was stopped after 200 generations, since it was found that most of the calculations were converged after less than 50 generations. A typical set of GA/LS runs for the host (S,S)-dimethyldiketopyridino-18-crown-6 **3** with methoxyalanine **G3** in methanolic solution with different sets of initial conditions is shown in fig 6.

4.5.2 Flexible Docking

Energy surfaces frequently show a number of deep minima that contribute significantly to the binding free energy of the system[16]. These minima often arise by side chain torsions or different binding modes of the ligand. If different modes are found, it has proven useful to classify the structures by energetic and structural parameters (e.g. rms values). However, the anchoring mechanism of the ammonium-crown interaction strictly enforces a deep minimum in the potential energy surface[19], such that the PES is considerably simplified. But, there is still a strong need for an appropriate sampling within the calculation of meaningful thermodynamic variables[16].

Energetic criteria were used to select 10 relevant minima from the rigid-body GA run including the global minimum. These structures were subjected to a flexible optimization in the RISM solvent by use of the Simplex algorithm (eq 4.3). During optimization, no limit of the maximum number of steps was given. The calculation was considered converged, when the criterion

$$|\Delta W(\mathbf{\Gamma}_{XY})| < 10^{-4} \text{ kcal mol}^{-1} \tag{4.15}$$

was met.

It is illustrative to calculate Boltzmann-factors γ_i associated with individual minima W_i, since they allow to estimate the statistical relevance of the subset:

$$\gamma_i = \frac{\exp(-W_i / kT)}{Z}, \tag{4.16}$$

where the partition function Z is approximated as

$$Z = \sum_i \exp(-W_i / kT). \tag{4.17}$$

The resulting Boltzmann-factors for the extrema of individual minima γ_{min} and γ_{max} are given in Table 1. It is found, that all factors γ_{max} are large, accounting for the

dominant contribution of the global minimum to the energy of the ensemble. Besides, the complexes **2-G1**, **4-G1**, **6-G1**, **3-G1** exhibit a small γ_{min}, supporting the assumption that the configurational space of the bonded complexes is well represented by the set of 10 minima. However, for **7-G2, 3-G2**, **3-G4**, **3-G5** we find γ_{min} still contributing markedly to the energy of the ensemble. In these cases, an enlarged set of more than 10 minima would probably improve the statistical relevance.

TABLE 1: **Resulting extrema of the Boltzmann factors of individual minima derived by the LGA/RISM calculations**.

	$\gamma_{R,\,min}$	$\gamma_{R,\,max}$	$\gamma_{S,\,min}$	$\gamma_{S,\,max}$
2-G1	0.0121	0.1890	0.0030	0.3194
7-G2	0.0440	0.1792	0.0070	0.2995
5-G1	0.0241	0.3016	0.0160	0.2096
4-G1	0.0004	0.5476	0.0014	0.2667
6-G1	0.0099	0.4271	0.0246	0.197
3-G1	0.0024	0.5090	0.0011	0.4557
3-G2	0.0377	0.1866	0.0066	0.2400
3-G3	0.0350	0.2431	0.0015	0.3229
3-G4	0.0083	0.1842	0.0326	0.2905
3-G5	0.0644	0.1373	0.0541	0.1767

4.6 Thermodynamics of chiral selection

Thermodynamic parameters have been determined from the RISM/LGA solutions. The expectation value of the individual host-guest interaction energy ΔU was calculated by a statistical average of ten minima W_i resulting from flexible mimimization:

$$\Delta U = \sum_i \gamma_i W_i \,, \tag{4.18}$$

The individual free energy of complexation is derived by

$$\Delta A = -k_B T \ln\langle Z \rangle . \tag{4.19}$$

The neglection of a pressure-volume term[48] yields

$$\Delta G \cong \Delta A . \tag{4.20}$$

Eqs 4.18 and 4.19 yield a crude approximation to the complexation entropy

$$T\Delta S = \Delta A - \Delta U \tag{4.21}$$

Chiral selections of diastereomeric complexes have been calculated in terms of the individual free energies of complexation (eq 4.20):

$$\Delta\Delta G_{R-S} = \Delta G_R - \Delta G_S \tag{4.22}$$

TABLE 2: Thermodynamic parameters of the diastereomeric host-guest complexes used in this study. All parameters are given in units of kcal mol^{-1}. $\Delta W_{R\text{-}S}^{rigid}$ was derived by the rigid predocking step, while $\Delta W_{R\text{-}S}$ is the difference of the global minima.

	$\Delta W_{R\text{-}S}^{rigid}$ / kcal mol^{-1}	$\Delta W_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta U_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta(TS_{R\text{-}S})$ / kcal mol^{-1}	$\Delta\Delta G_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta G_{exp}$ [12] / kcal mol^{-1}
2-G1	0.385	1.384	1.602	0.093	1.695	0.33
7-G2	0.748	1.307	1.205	0.407	1.611	0.19
5-G1	0.718	1.978	1.831	-0.068	1.763	0.17
4-G1	0.001	1.786	1.649	-0.289	1.360	0.58
6-G1	0.622	0.653	0.547	-0.353	0.194	0.36
3-G1	0.777	1.923	1.857	0.001	1.858	0.58
3-G2	-0.704	2.357	2.378	0.128	2.506	0.30
3-G3	0.351	0.283	0.285	0.166	0.451	0.33
3-G4	0.916	0.102	0.369	0.003	0.372	-
3-G5	0.385	0.619	0.738	0.031	0.769	-

The corresponding thermodynamic parameters are given in Table 2. The differences $\Delta\Delta(TS_{R\text{-}S})$ are found to be small, indicating a very similar binding mode of the R-guest and the S-guest. This might lead to a significant cancellation of errors in our free energy calculations.

It is found, that all (S,S)-hosts bind stronger to (R)-ammonium guests than to the corresponding (S)-guests. The calculated differences of the interaction energies $\Delta\Delta U_{R\text{-}S}$ and free energies $\Delta\Delta G_{R\text{-}S}$ are positive in sign. Obviously, the guest molecules with a configuration opposite to the configuration of the host are preferred in complexation. This primary result corresponds well with experimental results[12]. The absolute values of the chiral selections determined by our method agree reasonably well with experimental data.

However, the difference of the global energetic minima of the diastereomeric complexes $\Delta W_{R\text{-}S}$ do also account well for the chiral discrimination. This is not a coincidence, since the calculated Boltzmann factors suggest that the complexation energy surfaces exhibit predominant rather than shallow global minima. Most likely, a small number of minima play a key role in the complexation free energy surface. The global energetic minima of the rigid pre-docking step match the experimental data well. Within the set of complexes examined in this study, this is due to the fact that the minimum energy components in solution are structural similar to the complex structures, e.g. the flexibility of the complex partners is small and little distortion of the molecules is necessary in order to form complexes.

The chiral discrimination process is best understood in terms of the molecular structure. It is known to arise from the interaction of side chains of the diastereomeric guests with the asymmetric centers of the host. The selectivity of the system ammonium/pyridino-18-crown-6 is enhanced by strong hydrogen bonding and limited conformational flexibility. In fig 7, 10 relevant minima of the system **G1-3** resulting from the RISM/LGA solution are shown.[49,50] The complex structures of both the (R)-**G1-3** and the (S)-**G1-3** stereoisomer are found to be dominated by hydrogen bonding of the ammonium and aromatic stacking between the naphtyl group of **G1** and the pyridino group of the host. The methyl group in position 2 of the host acts as chiral selector by interaction with the naphtyl group of **G1**. By this means, aromatic stacking of (S)-**G1-3** is hindered more than that of (R)-**G1-3**, accounting for the larger rms values observed in (R)-**G1-3** ($0.113 < \text{rms} < 0.447$) in comparison with (S)-**G1-3**, where we find ($0.106 < \text{rms} < 0.226$).

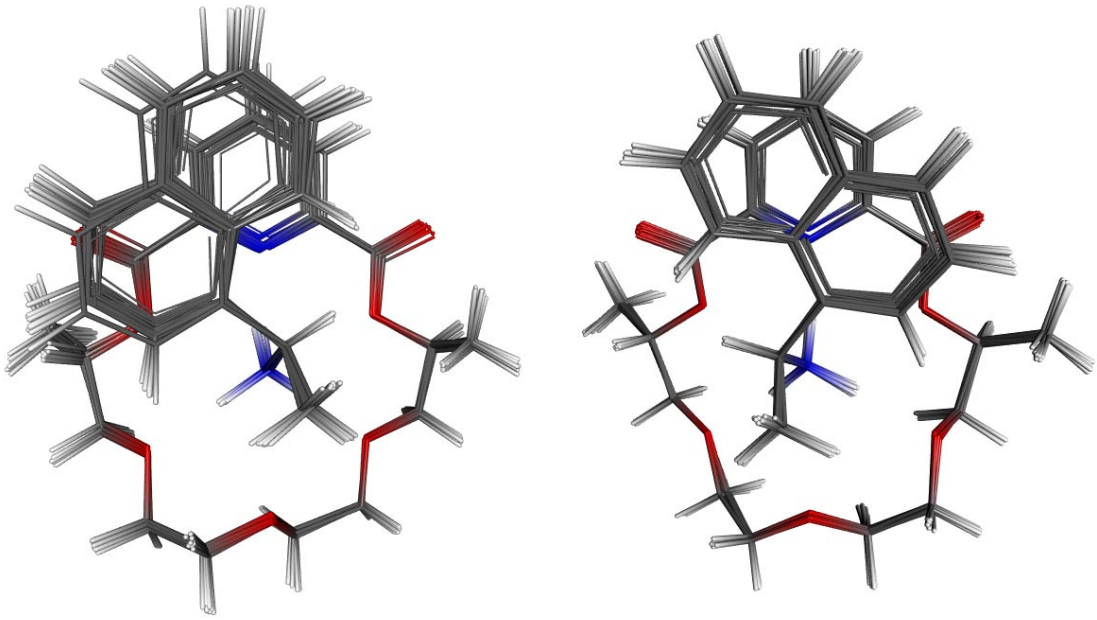

Figure 7: Recognition of guest (R)-**G1** (left hand side) and guest (S)-**G1** (right hand side) by host **3**. 10 relevant minima resulting from RISM/LGA calculations are shown. Chiral selection arises from interference of the guests' side chain with the stereochemical centers of the host.

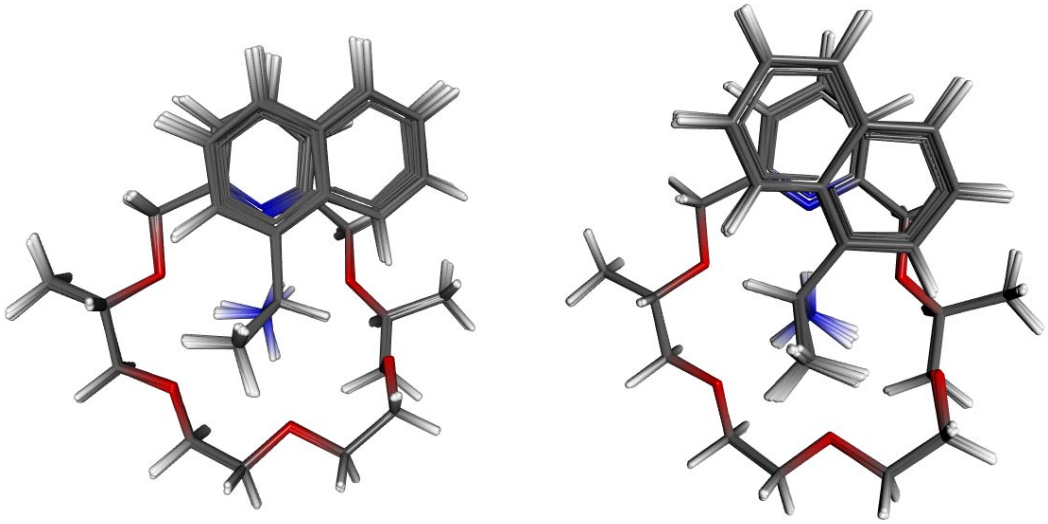

Figure 8: Recognition of guest (R)-**G1** (left) and guest (S)-**G1** (right) by host **2**. 10 lowest minima resulting from RISM/LGA calculations are shown. The conformation of the asymmetric centers of the host **2** differs from that of **3**, leading to weakening of selection.

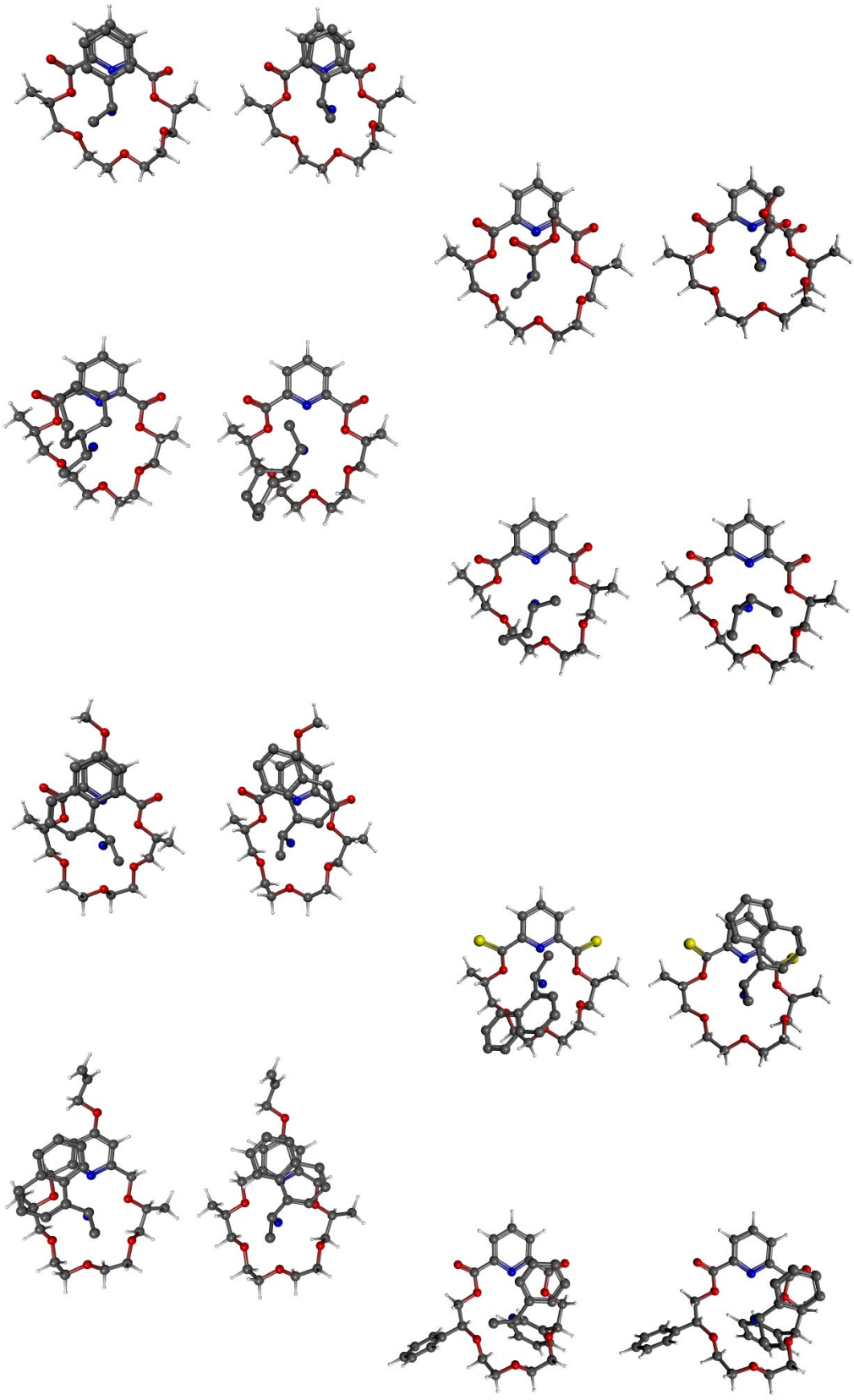

Figure 9: Diastereomeric pairs of the residue host-guest complexes examined in this study. From top to bottom: **3-G2**, **3-G3**, **3-G4**, **3-G5**, **4-G1**, **5-G1**, **6-G1**, **7-G2**.

It is illustrative to compare these results with the system **G1-2**, where the keto groups of the host have been replaced by methylene groups (fig 8). Here, the deviation of the stereoselective methyl groups from the horizontal plane of the crown ether is small, reducing the interaction with the guest and also the chiral selection. With **2**, we find a maximum rms value of 0.2 for the R-guest and 0.15 for the S-guest.

A complete overview of the chiral systems modeled is given in fig 9.

4.7 Conclusions

We have reported a new hybrid algorithm combining the reference interaction site model (RISM) integral equation theory for molecular liquids with a genetic algorithm for molecular docking in solution. The GA has been used in connection with a local search technique to scan the topography of the solvation potential energy surface most effectively.

The algorithm has been applied to perform docking with host-guest complexes of (S,S)-dimethyldiketopyridino-18-crown-6 derivatives with chiral organic ammonium cations in methanolic solution. The docking process was performed in two steps. First, a rigid body pre-docking step was run with the RISM/LGA algorithm to result a number of meaningful local minima. Second, these minima were subjected to flexible optimization with the TRIPOS all atom force field. From these solutions of the RISM/LGA algorithm, thermodynamic parameters of the host-guest complexes have been determined by statistical averaging, such as $\Delta\Delta G_{\text{R-S}}$, $\Delta\Delta U_{\text{R-S}}$ and $\Delta\Delta(TS_{\text{R-S}})$. Chiral selections of the diastereomeric host-guest complexes were calculated in terms of their individual free energies of complexation. It was found that the algorithm is throughout capable to predict the sign of the chiral discrimination within host-guest complexes. The calculated differential free energies of selection $\Delta\Delta G_{\text{R-S}}$ were found to be in good agreement with experimental measurements.

The mechanisms of chiral discrimination have been studied by means of the minimum-energy structures derived by the RISM/LGA algorithm. Indications have been found for steric interactions of the naphtyl group with the asymmetric centers in the (S)-**G1-3** complex. These interactions were found to be smaller in the isomeric (R)-**G1-3** complex.

4.8 References

(1) Lehn, J.-M. Supramolecular Chemistry: Concepts and Perspectives, VCH, Weinheim **1995**.

(2) McDonald, N.A.; Duffy,E.M., Jorgensen, W.L. *J. Am. Chem. Soc.* **1998**, *120*, 5104.

(3) Wang, J.; Kollman, P.A. *J. Am. Chem. Soc.* **1998**, *120*, 11106.

(4) Reinhardt, W.P.; Miller, M.A., Amon, L.M. *Acc. Chem. Res.* **2001**, *34*, 607.

(5) Schlick, T.; Skeel, R.; Brünger, A.; Kalé, L.; Board Jr.; J.A., Hermans, J.; Schulten, K. *J. Comp. Phys*. **1999**, *151*, 9.

(6) Izrailev, S.; Crofts, A.R.; Berry, E.A.; Schulten, K., *Biophys. J*. **1999**, *77*, 1753.

(7) Hermansson, J. *J. Chromatogr.* **1983**, *269*, 71.

(8) Ward T.J.; Armstrong, D.W. *J. Liq. Chromatography* **1986**, *9*, 407.

(9) Armstrong, D.W.; Tang, Y.; Chen, S.; Zhou, Y.; Bagwill, C.; Chen, J.-R. *Anal. Chem*. **1994**, *66*, 1473.

(10) Hedges, A.R. *Chem. Rev*. **1998**, *98*, 2035.

(11) Vespalec, R.; Boek, P. *Chem. Rev*. **2000**, *100*, 3715.

(12) Izatt, R.M. *Chem. Rev*. **1997**, *97*, 3313.

(13) http://www.chiraltech.com/new/navortispress.htm

(14) Horváth, G.; Huszthy, P.; Szarvas, S.; Szókán, G.; Redd, J.T.; Bradshaw, J.S.; Izatt, R.M. *Ind. Eng. Chem. Res.* **2000**, *39*, 3576.

(15) Salem, L.; Chapuisat, X.; Segal, G.; Hiberty, P.C.; Minot, C.; Leforestier, C.; Sautet, P. *J. Am. Chem. Soc.* **1987**, *109*, 2887.

(16) Lipkowitz, K.B. *Acc. Chem. Res.* **2000**, *33*, 555.

(17) Rudolph, J. Selektive Erkennung von Molekülen und stereotypen funktionellen Gruppen, Ph.D. thesis, Bochum **1995**.

(18) Schmidt, K.F.; Kast, S.M. *J. Phys. Chem. B*, submitted for publication.

(19) Schmidt, K.F.; Kast, S.M., Brickmann, J. in preparation.

(20) Chandler, D.; Andersen, H.C. *J. Chem. Phys.* **1972**, *57*, 1930.

(21) Hirata, F.; Rossky, P.J. *Chem. Phys. Lett.* **1981**, *83*, 329.

(22) Lee, P.H.; Maggiora, G.M. *J. Phys. Chem.* **1993**, *97*, 10175.

(23) Singer, S.J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

(24) Karplus, M. CHARMM Documentation, The CHARMM Development Project, Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

(25) Craig, D.P.; Elsum, I.R. *Chem. Phys.* **1982**, *73*, 349.

(26) Cann, N.M.; Das, B. *J. Chem. Phys.* **2000**, *113*, 2369.

(27) Holland J.H. Adaptation in natural and artificial system, Ann Arbor, The University of Michigan Press, **1975**.

(28) Heistermann, J. Genetische Algorithmen - Theorie und Praxis evolutionärer Optimierung, B.G. Teubner Verlagsgesellschaft, Stuttgart **1994**.

(29) Goldberg D. Genetic Algorithms, Addison Wesley, 1988.

(30) Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. *J. Comput. Chem.* **1998**, *19*, 1639.

(31) Dengiz, B.; Altiparmak F.; Smith, A.E. IEEE Transactions on Evolutionary Computation **1997**, *1*, 179.

(32) Freisleben, B.; Merz, P. Proc. IEEE International Conference on Evolutionary Computation, IEEE, World Congress on Computational Intelligence, **1996**, 616.

(33) Gottlieb, J. Evolutionary Algorithms for Constrained Optimization Problems, PhD thesis, Department of Computer Science, Technical University of Clausthal, Germany **1999**.

(34) SYBYL 6.6, *Tripos Inc.*, 1699 South Hanley Road, St. Louis, Missouri, 63144 USA.

(35) Clark, M.; Cramer III, R.D.; van Opdenbosch, N. *J. Comp. Chem.* **1989**, *10*, 982.

(36) Numerical Recipes in C: THE ART OF SCIENTIFIC COMPUTING, Cambridge University Press, **1988**, 408.

(37) Pettitt, M.B.; Rossky, P.J. *J. Chem. Phys.* **1986**, *84*, 5836.

(38) Hansen, J.P.; McDonald, I.R. Theory of simple liquids, Academic Press, London **1980**.

(39) Alcaro, S.; Gasparrini, F.; Incani, O.; Mecucci, S.; Misiti, D.; Pierini, M.; Villani, C. *J. Comput. Chem.* **2000**, *21*, 515.

(40) Davidson, R.B.; Bradshaw, J.S.; Jones, B.A.; Dalley, N.K.; Christensen, J.J.; Izatt, R.M.; Morin, F.G.; Grant, D.M. *J. Org. Chem.* **1984**, *49*, 353.

(41) Dearden, D.V.; Dejsupa, C.; Liang, Y.; Bradshaw, J.S.; Izatt, R.M. *J. Am. Chem. Soc.* **1997**, *119*, 353.

(42) Lee, O.-S.; Hwang, S.; Chung, D.S. *Supramol. Chem.* **2001**, *12*, 255.

(43) Jorgensen, W.L. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(44) Jorgensen, W.L. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(45) Kovalenko, A.; Ten-no, S.; Hirata, F. *J. Comp. Chem.* **1999**, *20*, 928.

(46) Talman, J.D. *J. Comput. Phys.* **1978**, *29*, 35.

(47) Carroll, D.L. program gafortran (version 1.7), University of Illinois, 306 Talbot Lab, 104 S. Wright St., Urbana, IL 61801, **1998**.

(48) Gilson, M.K.; Given, J.A.; Bush, B.L.; McCammon, J.A. *Biophys. J.* **1997**, *72*, 1047.

(49) Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C.D. *Interactive Visualization of Molecular Scenarios with the MOLCAD/SYBYL Package*. In: Bowie, J.E.; Ed.; Data Visualization in Molecular Science "Tools for Insight and Innovation", Addison-Wesley Publishing Company Inc., Reading, Mass. **1995**.

(50) Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G. *Molecular Models: Visualization.* In: Schleyer, P.v.R.; Allinger, N.L.; Clark, T.; Gasteiger, J.; Kollman, P.A.; Schaefer III, H.F.; Schreiner, P.R.; Eds., "The Encyclopedia of Computational Chemistry", John Wiley & Sons: Chichester, **1998**.

5 Summary and Conclusions

In the last decades, tremendous progress has been made in the computer-assisted study of complex heterogenous systems. Detailed free energy simulation of these system on the atomic level has become feasible, but suffers from the large computational effort which is usually spent in sampling the configurational space of the solvent. The calculation times can be drastically reduced using statistical solvent models, such as the RISM integral equation theory, which offers a traceable route to free energies of solvation.

It has been the central intention of this thesis to combine RISM integral equation theory with well-established computational techniques to explore free energy surfaces (FES) of supramolecular association by means of hybrid approaches. The following three approaches have been developed for that purpose:

(a) A computer program has been developed, that combines a RISM integral equation representation of the solvent with the optimization of a fully flexible all-atom force field representation of molecular complexes. The determination of molecular structures in solution takes place under more realistic conditions than provided by vacuum force fields due to accounting for the solvation free energy of the complex. Besides, optimization of complex structures with distance constraints was found to be a fast and easy way to characterize valleys on the FES of complexation. If leading to the global minimum, those valleys can be considered as minimum-energy-pathways related to the reaction mechanism on the RISM-derived FES. By this means, the complex formation of alkali metal ions and organic ammonium ions with 15-crown-5 and 18-crown-6 was studied in water, methanol and acetonitrile.

(b) The traditional atomistic Metropolis Monte-Carlo simulation of molecular complexation has been combined with an integral equation representation of the solvent phase. The technique of free energy perturbation has been used to derive potentials of mean force of intermolecular associations from which the computation of the free energy of complexation is possible. Using the hybrid RISM/MC method, extensive studies have been performed concerning the

complexation thermodynamics of alkali metal ions, guanidinium derivatives and organic ammonium ions with 15-crown-5 and 18-crown-6 in water, methanol and acetonitrile. Although the 1D-RISM equations within the HNC approximation suffer from well-known limitations, the results obtained were found to agree reasonably well with experimental findings. Within each series, the selectivity of the host was shown to be consistent with the experimental data.

(c) A genetic algorithm has been combined with integral equation theory to perform docking studies. This allows for global optimization of the complexation FES. The GA has been used in connection with a local search technique to scan the topography of the solvation potential energy surface most effectively. The RISM/LGA method has been applied to examine diastereomeric host-guest complexes in methanolic solution. Chiral selections of the diastereomeric host-guest complexes have been calculated in terms of their individual free energies of complexation and were found to be in good agreement with experimental measurements. Specific discrimination mechanisms have been examined in terms of the minimum-energy structures derived by the RISM/LGA method. The rms values of the diastereomeric complexes provided evidence for selective steric interactions of the guest's side chains with the asymmetric centers of the host.

It has been demonstrated in this work, that problem-oriented hybrid algorithms combining simulation techniques and integral equation theory are very effective tools in the computational treatment of complexation thermodynamics and solution structure. The efficiency of free-energy simulations is improved drastically by use of RISM. Within the examples presented, the calculation times are reduced by an order of magnitude as compared to molecular dynamics or Monte-Carlo simulations. The accuracy of the results is still convincing, even if double differences of thermodynamic variables (such as chiral selection data) are considered.

However, the application of RISM is presently limited to those cases where the solvent structure and dynamics is of minor importance. The use of RISM as an effective potential in molecular dynamics simulations seems a promising future application. It might probably help in the clarification and even compensation of methodical errors in both simulation techniques and integral equation theory [S.M. Kast, private communication].

6 Zusammenfassung

6.1 Einleitung

Die Chemie supramolekularer Verbindungen ist in den letzten Jahren auf wachsendes Interesse in der Forschung gestoßen.[1] Anfänglich dienten einfache supramolekulare Komplexe in Lösung als Studien- und Referenzmodelle zur Modellierung der Eigenschaften komplexer Biomoleküle, zu denen oft kein unmittelbarer Zugang bestand. Eine wichtige Eigenschaft dieser Substanzklasse besteht in der hohen *Selektivität* vieler *Wirtsmoleküle* für einen bestimmten *Gast*[2]. Die dabei beobachteten Bindungsmoden sind ausgesprochen vielseitig. Supramolekulare Verbindungen können durch elektrostatische oder hydrophobe Wechselwirkungen, Wasserstoffbrückenbindungen oder Formkomplementarität stabilisiert werden. Eine Sonderstellung nimmt dabei die Gruppe der Metallkationen-Komplexe ein, die oft Einschlußverbindungen bilden.[3]

In der synthetischen und technischen Chemie werden diese Eigenschaften zur Stofftrennung und selektiven Katalyse genutzt.[4] Besonders die pharmakologische Wirkstoffentwicklung erzwingt einen sehr hohen Reinheitsgrad der Edukte und Produkte zur Minimierung unerwünschter Nebenwirkungen bei Präparaten. Eine Sonderstellung nehmen dabei Verfahren zur Erzeugung enantiomerenreiner Produkte ein. Von wenigen Ausnahmen abgesehen,[5] können reine Stereoisomere in der Synthese und in Trennverfahren ausschließlich über die Wechselwirkung mit einem chiralen Selektor dargestellt werden.[6] In den letzten Jahren zeigten zahlreiche Arbeiten auf den Gebieten der biomimetischen und asymmetrischen Katalyse[7,8,9] innovative Möglichkeiten zur technischen Nutzung der *stereochemisch selektiven, molekularen Erkennung*, deren Bedeutung in der Proteinchemie seit langem bekannt ist[10], und der chromatographischen Trennung racemischer Gemische.[11,12,13,14] Besondere Erfolge bei der chiralen Selektion von Ammoniumverbindungen und anderen Substraten konnten insbesondere mit chiralen Kronenethern[15,16] und Cyclodextrinderivaten[17,18] erzielt werden. Derartige Komplexe, bei denen neben anderen Faktoren die Formkomplementarität eine große Rolle spielt, werden in der Literatur häufig kurz als *Wirt-Gast-Systeme* bezeichnet.

Die experimentell zugänglichen Kenngrößen supramolekularer Verbindungen in Lösung beinhalten strukturelle Eigenschaften und thermodynamische Größen. So läßt sich die thermodynamische Stabilität einer Verbindung durch die *Freie Bindungsenthalpie* ΔG bestimmen. Analog dazu kann die relative Stabilität der Assoziate im Sinne der Gleichgewichtsthermodynamik mit der Differenz der Freien Bindungsenthalpien $\Delta\Delta G$ der isomeren Gastmoleküle mit dem Wirtsmolekül identifiziert werden. Hier liegt eine wesentliche Schnittstelle mit dem Computerexperiment, welches durch geeignete Modellierung des Systems eine Berechnung und Voraussage der Kenngrößen leisten kann. Die schnelle und gleichzeitig hinreichend genaue Berechnung thermodynamischer Eigenschaften ist die Basis zur Untersuchung *supramolekularer Verbindungen* in Lösung mit Hilfe von Computersimulationsmethoden. Während molekulare Simulationen auf der Basis atomistischer Modelle oft große Rechenkapazitäten erfordern, kann der Aufwand durch den Einsatz von statistisch-mechanischen Integralgleichungsansätzen stark reduziert werden. Im Rahmen der vorliegenden Arbeit sollten daher Integralgleichungsmethoden auf ihre Eignung zur Berechnung der Thermodynamik von Komplexbildungen von Kronenethern in Lösung untersucht werden. Insbesondere sollte dabei auch eine Computersimulationsmethode zur quantitativen Bestimmung von Selektions-Effekten bei der Bildung chiraler Kronenether-Komplexe entwickelt werden.

Prinzipiell können zwei Strategien zur Berechnung Freier Bindungsenthalpien bei Komplexbildungen in Lösung verfolgt werden, die sich in ihrem Zugang zu der Größe ΔG unterscheiden:

Einerseits lassen sich empirische Verfahren formulieren, die häufig auf der Extrapolation thermodynamischer Konzepte auf molekulare Dimensionen beruhen, zum anderen kann man versuchen, ΔG auf der Basis atomar mikroskopischer Modelle zu berechnen. Bekannte Verfahren der ersten Kategorie umfassen Untersuchungen auf der Basis molekularer Oberflächenkomplementarität,[19,20,21] sowie der empirischen quantitativen Struktur-Wirkungs-Beziehung (*QSAR*/*QSPR*)[22,23,24] und dreidimensionaler Molekülwechselwirkungsfelder (*CoMFA*).[25] Diesen Verfahren ist gemeinsam, daß die intermolekularen Wechselwirkungen auf der Basis einer empirischen Zielfunktion evaluiert werden, deren Parametrisierungsstrategie oft der jeweiligen Problemstellung angepaßt werden muß. Die Kombination empirischer Zielfunktionen mit evolutionären Algorithmen (*EA*),[26,27] insbesondere genetischen

Algorithmen (*GA*),[28,29,30] hat zu der Entwicklung erfolgreicher *Docking-Methoden*[31,32,33,34] geführt.

Verfahren der zweiten Kategorie stützen sich zumeist auf Simulationsrechnungen. Molekulardynamische Untersuchungen[35] erlauben es, mittels thermodynamischer Integration[36] oder thermodynamischer Störungstheorie[37,38,39] Freie Komplexbildungsenthalpien in Lösung zu bestimmen.[40] Hierbei werden „samples", d.h. Stichproben von molekularen Konfigurationen statistisch erzeugt, und hieraus die thermodynamische Variable durch Mittelung berechnet. Auch die statistische Monte-Carlo-Simulationsmethode (*MC*) hat bereits breite Anwendung zur Bestimmung von Komplexierungsenthalpien gefunden.[41,42] Die Qualität der Resultate wird dabei von dem verwendeten Kraftfeld und der Simulationszeit, beziehungsweise der Anzahl der erzeugten Konfigurationen, mitbestimmt. Der hierfür benötigte Aufwand ist nicht zuletzt durch den Einbezug der zahlreichen Freiheitsgrade des Lösungsmittels bisweilen beträchtlich.[43]

Eine erfolgversprechende Strategie zur Reduktion des Rechenzeitbedarfs im Rahmen atomistischer Simulationsverfahren bei gleichzeitigem Erhalt der Allgemeingültigkeit und brauchbarer Vorhersagekraft besteht in der Entwicklung hybrider Verfahren. Es bietet sich an, die Anzahl der Lösungsmittelkonfigurationen durch ein statistisches Verfahren zu approximieren. Nichtlineare, statistisch-mechanische Integralgleichungsmodelle[44] (*reference interaction site model*, RISM)[45] zur Beschreibung der Solvensdichte über radiale Verteilungsfunktionen sind dazu gut geeignet. Integralgleichungsmethoden ermöglichen eine schnelle Berechnung der Orts- und Orientierungskorrelationsfunktionen zwischen Molekülen auf der Basis der molekularen Struktur und eines geeigneten Kraftfeldes. Ausgehend von den Verteilungsfunktionen können thermodynamische Eigenschaften der Lösung, wie z.B. Freie Lösungsenthalpien, in ausreichender Genauigkeit[46] berechnet werden. Dieser Ansatz wurde in der vorliegenden Arbeit verfolgt. Ein erster Schritt hierzu besteht in der Berechnung von Energie-Hyperflächen der Komplexbildung in Lösung auf der Basis der Integralgleichungsmethode RISM. Dazu werden die Integralgleichungen mit einer geeigneten approximativen closure Gleichung kombiniert. Zur Berechnung thermodynamischer Variablen hat sich die hypernetted chain-Gleichung (*HNC*) [47] bewährt.

6.2 Energie-Hyperflächen

Die Lösung der RISM Grundgleichungen auf der Basis der hypernetted chain Gleichung ermöglicht die Angabe eines konfigurationsabhängigen Ausdrucks für die freie Solvatationsenthalpie $\Delta\mu_A$ eines Moleküls in unendlicher Verdünnung.[48] Mit der, ebenfalls konfigurationsabhängigen, Vakuumenergie E^{vac} kann die Energie-Hyperfläche des Moleküls A berechnet werden:

$$E_A^{tot}(\mathbf{\Gamma}_A) = E_A^{vac}(\mathbf{\Gamma}_A) + \Delta\mu_A(\mathbf{\Gamma}_A) \tag{5.1}$$

Dabei steht E^{tot} für die Gesamtenergie einer Molekülkonformation mit den Koordinaten $\mathbf{\Gamma}_A$. Während intramolekulare Wechselwirkungen im Rahmen eines Kraftfeldansatzes[47] gut beschrieben werden können, ergibt sich bei intermolekularen Wechselwirkungen in Lösung eine kompliziertere Situation. Die Freie Komplexierungsenthalpie für eine gegebene Komplexstruktur in Lösung kann auf diesem Weg nur unter dem Einsatz aufwendiger Methoden berechnet werden. Alternativ kann diese Größe auf indirektem Weg über einen thermodynamischen Kreisprozeß ermittelt werden. Die Reaktionsenergie-Hyperfläche der Komplexbildungsreaktion

A + B → [AB]

kann auf der Basis von Gleichung 5.2 aus der Summe der Komponenten (Gl. 5.1) berechnet werden. Dazu wird entsprechend Abb. 1 ein thermodynamischer Kreisprozeß für die Komplexbildung formuliert.

$$\begin{aligned} W(\mathbf{\Gamma}_{AB}) = {} & \Delta V_{AB}(\mathbf{\Gamma}_{AB}) + E_{AB}^{vac}(\mathbf{\Gamma}_{AB}) + \Delta\mu_{AB}(\mathbf{\Gamma}_{AB}) \\ & - \left[E_A^{vac}(\mathbf{\Gamma}_A) + \Delta\mu_A(\mathbf{\Gamma}_A)\right] - \left[E_B^{vac}(\mathbf{\Gamma}_B) + \Delta\mu_B(\mathbf{\Gamma}_B)\right] \end{aligned} \tag{2}$$

$W(\mathbf{\Gamma}_{AB})$ steht für das orientierungsabhängige potential-of-mean-force (*PMF*) des Komplexes [AB]. Dabei steht ΔV_{AB} für die Reaktionsenergie von [AB] im Vakuum, $\mathbf{\Gamma}_A$ für die intramolekularen Freiheitsgrade von A und $\mathbf{\Gamma}_B$ für die intramolekularen Freiheitsgrade von B.

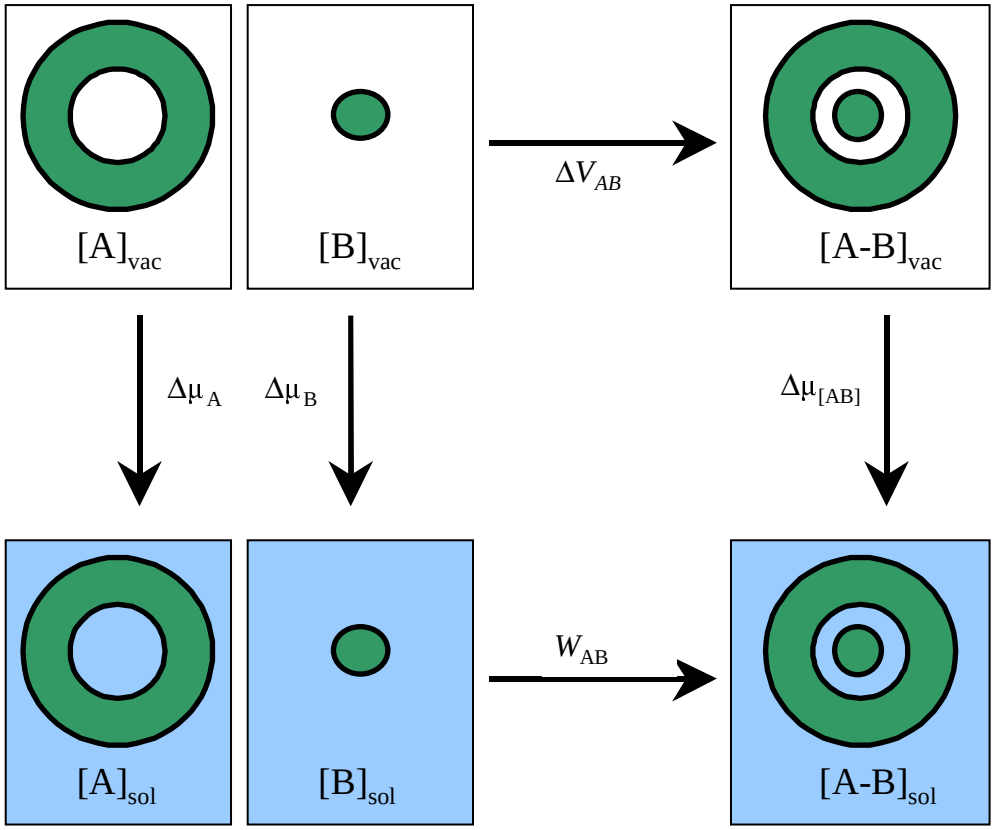

Abbildung 1: Konstruktion eines thermodynamischen Kreisprozesses für die Komplexierungsreaktion A + B → [AB] in Lösung. $\Delta\mu_A$, $\Delta\mu_B$ und $\Delta\mu_{[AB]}$ bezeichnen die freien Lösungsenthalpien von A, B und dem Komplex [AB] in einer gegebenen Konformation. ΔV_{AB} steht für die Reaktionsenergie in Vakuum.

Die Dimension der Hyperfläche in Gleichung 5.2 ist von der Anzahl molekularer Freiheitsgrade bestimmt. Im Falle der Wechselwirkung zweier starrer Moleküle reduziert sich diese auf sechs – den Abstand der Massenschwerpunkte und fünf unabhängige Rotationswinkel. Diese Freiheitsgrade der intermolekularen Orientierung, sind in der Abbildung 2 schematisch dargestellt. Sie setzen sich zusammen aus dem Abstand der Reaktionspartner *r*, drei Winkeln zur Beschreibung der Orientierung von A im Koordinatensystem von B {Θ, Φ, Ψ}, und zwei weiteren Winkeln {θ, φ}, welche die Lage des Abstandsvektors von A zu B im Raum angeben.

$$W^{\text{rigid}}(\mathbf{\Gamma}_{AB}) = \Delta V_{AB}(\mathbf{\Gamma}_{AB}) + \Delta\mu_{AB}(\mathbf{\Gamma}_{AB}) - \Delta\mu_A(\mathbf{\Gamma}_A) - \Delta\mu_B(\mathbf{\Gamma}_B) \tag{5.3}$$

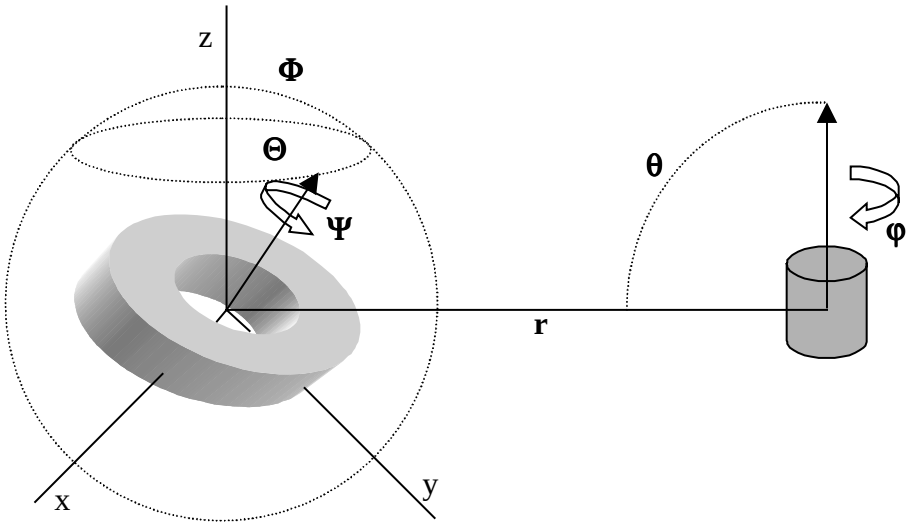

Abbildung 2: Polarkoordinatensystem zur Beschreibung der intermolekularen Freiheitsgrade in einem Wirt-Gast-System. Der Abstand der Reaktionspartner wird mit r bezeichnet. Drei Winkel {Θ, Φ, Ψ} werden zur Beschreibung der Orientierung von A im Koordinatensystem von B benötigt und zwei weitere Winkel {θ, φ} zur Beschreibung der Lage des Abstandsvektors von A zu B im Raum.

Bei der Komplexierung eines Metallkations durch einen starren Kronenether reduziert sich die Dimension der Hyperfläche auf drei, da in diesem Fall die Rotation des kugelsymmetrischen Kations für die Komplexierung nicht von Bedeutung ist. In solchen einfachen Fällen läßt sich eine partielle Energie-Hyperfläche der Komplexierung in Form eines Rastergitters punktweise konstruieren. Diese Methode wurde zur Untersuchung der Energie-Hyperflächen im System Metallkation/18-Krone-6 eingesetzt. In Kapitel 2.4 werden Schnitte durch die Freie-Energie-Hyperflächen der Metallkationen Na^+, K^+, Rb^+ und Cs^+ mit 18-Krone-6 als Komplexbildner in Methanol und Acetonitril angegeben. Die lokale Interferenz der Solvensschalen führt zur Ausbildung niedriger Barrieren, die durch thermische Anregung überwunden werden können. Die Graphen erlauben die Abschätzung der relativen Barrierenhöhen in den Lösungsmitteln für einen kollinearen Stoß der Komplexbildner zu 0.5 *kT* bis maximal 2 *kT*.

Im allgemeinen ist die in der Energie-Hyperfläche erhaltene Information über den gesamten Phasenraum zu umfangreich und läßt keinen Vergleich mit experimentell zugänglichen Größen zu. Gleichung 5.2 eröffnet jedoch die Möglichkeit, mit Optimiermethoden und Simulationsmethoden Erwartungswerte meßbarer energetischer Größen zu berechnen. Dies wird im Folgenden für den Fall der lokalen

Optimierung auf der Basis eines Simplex-Downhill-Algorithmus,[49] der Monte-Carlo-Simulationsmethode[37] und der globalen Optimierung durch einen genetischen Algorithmus[50] gezeigt.

6.3 Lokale Optimierung

Die lokale Optimierung von Komplexstrukturen dient zur Auffindung lokaler Minima auf der Freie-Energie-Hyperfläche, die energetisch bevorzugten Komplexstrukturen entsprechen. In gängigen Kraftfeldansätzen wird das Lösungsmittel oft nicht berücksichtigt, obwohl es die Konformation eines Komplexes beeinflussen kann. Diese Unzulänglichkeit kann mit der Integralgleichungsmethode vermieden werden. Die Lösung der RISM-Gleichungen führt zu einem konformationsabhängigen[48,51] Ausdruck für die freie Lösungsenthalpie, so daß Solvenseffekte in die Strukturoptimierung einbezogen werden können.

Zur schnellen und zuverlässigen Bestimmung lokaler Minima wurde ein Simplex-Downhill-Algorithmus[49] mit der Integralgleichungsmethode kombiniert, und zur Minimierung der Gesamtenergie von Komplexen (Gl. 5.2) eingesetzt. Der resultierende Algorithmus ermöglicht die lokale Optimierung bei selektiver Freigabe molekularer Freiheitsgrade, so daß auch starre oder teilstarre Systeme behandelt werden können. Die lokale Optimierung teilstarrer Systeme wurde im Rahmen der Arbeit zur Ermittlung möglicher Reaktionspfade der Komplexbildung eingesetzt, während die lokale Optimierung vollflexibler Systeme zur Bestimmung von Komplexgeometrien beim Docking chiraler Wirt-Gast-Systeme diente.

Im Falle der Komplexierung der Metallkationen Na^+, K^+, Rb^+ und Cs^+ mit 18-Krone-6 als Komplexbildner in Methanol und Acetonitril erfolgte die Bestimmung energetisch bevorzugter Reaktionspfade auf der Basis lokaler Optimierung und, darauf folgend, farbcodierter Visualisierung.[52,53] Dazu wurden in Abständen von 0.1 Å Kugelschalen um das Wirtsmolekül gelegt und Metallkationen nach dem Zufallsprinzip auf den Kugelschalen positioniert. Eine sich anschließende Energieminimierung im Raum der Polarkoordinaten $\{\Theta, \Phi\}$ lieferte bevorzugte radiale Positionen der Metallkationen rund um den Kronenether, welche sich im Nahbereich zu durchgängigen Pfaden verbinden lassen. Diese sind in Abb. 3 dargestellt. Dabei findet sich die Symmetriegruppe D_{3h} des Kronenethers in Form dreier rotationsäquivalenter Pfade mit horizontaler Spiegelebene auch in der Energie-Hyperfläche wieder. Entsprechend den aus der Literatur bekannten Ergebnissen bei der Kom-

plexierung von K^+ in Wasser[54] ist die kollineare Reaktion gegenüber dem seitlichen Angriff energetisch benachteiligt.

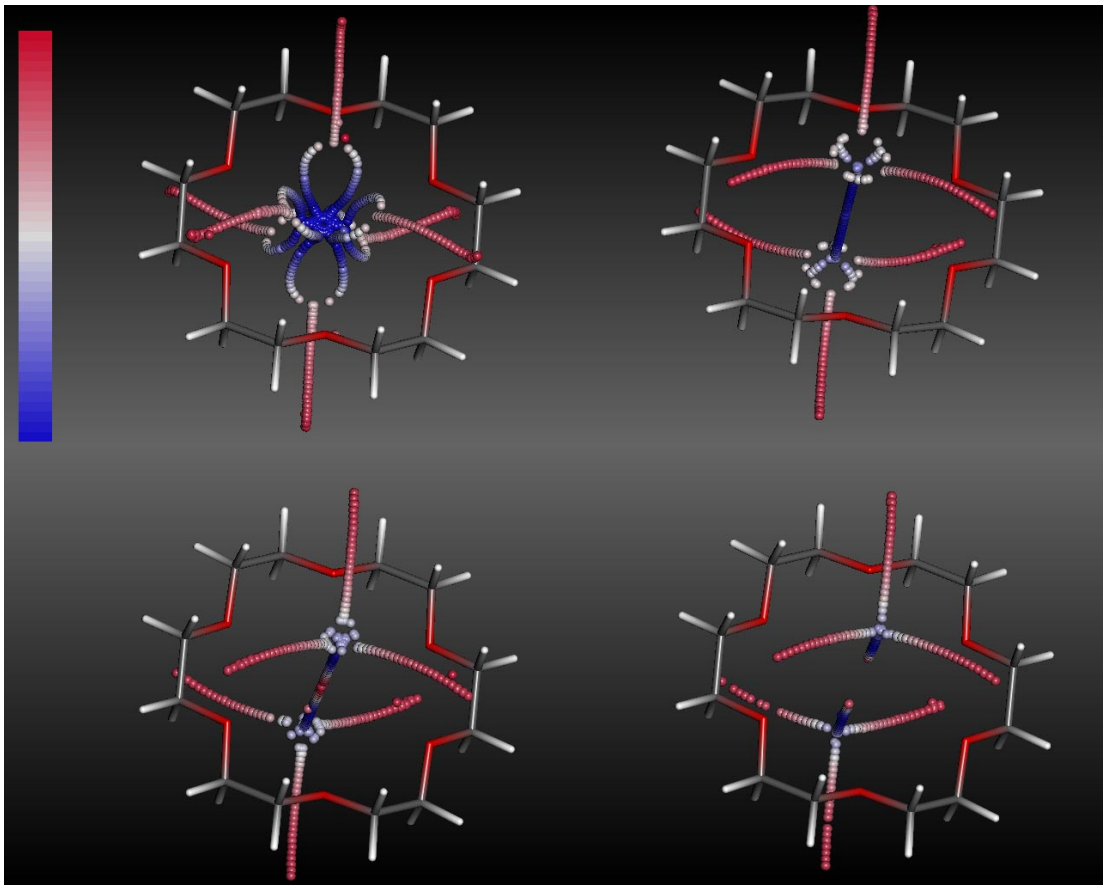

Abbildung 3: Farbcodierte Visualisierung bevorzugter radialer Positionen der Metallkationen Na^+ (obere Reihe, links), K^+; sowie Rb^+ (untere Reihe, links), Cs^+, rund um 18-Krone-6 im Bereich [0;7] Å. Die Position der Ionen ist durch Kugeln symbolisiert. Der Farbkeil verläuft von rot (geringe Bindungsenergie) nach blau (maximale Bindungsenergie).

Dieses Prinzip konnte auch bei der Komplexierungsreaktion des Methylammonium-Ions mit 18-Krone-6 in Wasser und Methanol beobachtet werden. Bei einer detaillierten Untersuchung der Energie-Hyperfläche dieses Systems wurden ebenfalls Minimum-Energie-Pfade gefunden, die eine C_3-Symmetrie bezüglich der Hauptachse des Kronenethers aufweisen. Diese Pfade sind in Wasser aufgrund der dielektrischen Eigenschaften von kurzer Reichweite. Es zeigte sich, daß hier eine schrittweise Bildung der drei bindenden Wasserstoffbrücken begünstigt ist, da im hydratisierten Kronenether ein verbrückendes Lösungsmittelmolekül die Struktur

stabilisiert und eine kollineare Reaktion verhindert. Die Pfade beginnen mit der Ausbildung eines locker gebundenen, einfach verbrückten Komplexes. Mit zunehmender Annäherung der Komplexpartner sinkt die Gesamtenergie des Systems, bis die Hauptachse des Gastmoleküls auf der C_3-Achse des Kronenethers zu liegen kommt.

Der Fall der vollflexiblen Optimierung von Wirt-Gast-Komplexen wurde zur computergestützten Modellierung von Komplexpartnern eingesetzt und findet im Rahmen der folgenden Unterkapitel weitere Erwähnung.

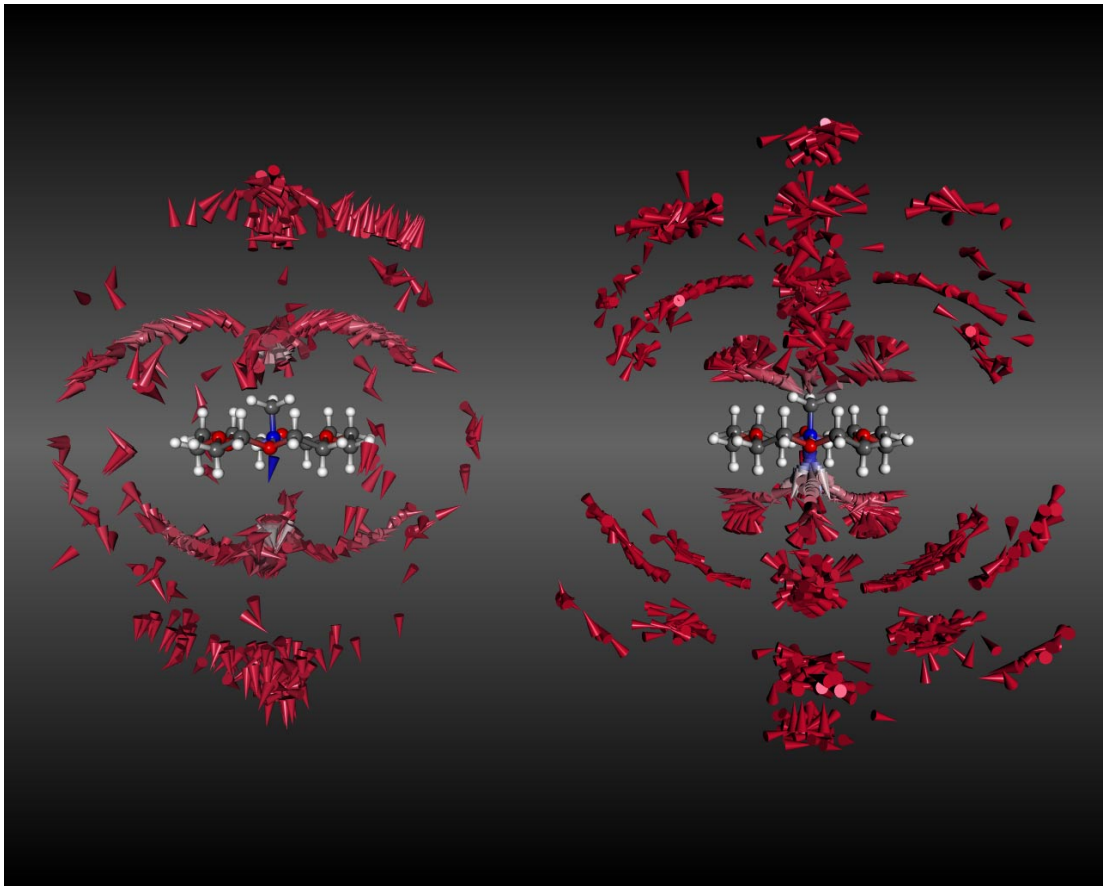

Abbildung 4: Farbcodierte Visualisierung bevorzugter Positionen von $H_3CNH_3^+$ um 18-Krone-6 im Bereich [0;15] Å. Die Kegelstümpfe bezeichnen die Lage der $(N)H_3$-Ebene, die Kegelspitzen die Positionen der Methylgruppe $C(H_3)$. Die Pfade in Wasser (links) sind von geringerer Reichweite als in Methanol (rechts). Der Farbkeil verläuft von rot (geringe Bindungsenergie) nach blau (maximale Bindungsenergie).

6.4 Thermodynamik bei Komplexbildungen

Schwach gebundene Systeme (Komplexe) weisen meist eine Vielzahl unterschiedlicher intermolekularer Konfigurationen auf. Somit ist der zur Bindung beitragende Konformationsraum groß. Zur Ermittlung Freier Bindungsenthalpien ist in diesem Fall die Methode des statistisches *Samplings* der intermolekularen Freien Energiefläche gut geeignet. Diese Technik ist im Rahmen der Molekulardynamik und Monte-Carlo-Methodik bereits etabliert, jedoch stets mit einem hohen numerischen Aufwand verbunden, der im wesentlichen auf die explizite Behandlung von Solvensmolekülen zurückgeht. Durch Kombination von Monte-Carlo- und Integralgleichungsmethode kann dieser Aufwand maßgeblich reduziert werden.

Zum schnellen *Sampling* Freier Energieflächen in Lösung wurde ein derartiges hybrides Verfahren entwickelt. Dabei wurde die Solvensphase durch die Lösung der zugehörigen Integralgleichung modelliert und eine Statistik über die Orientierungsfreiheitsgrade von Wirt und Gast auf der Basis der Metropolis-Monte-Carlo-Methode[55] durchgeführt. Das Verfahren wurde auf die Komplexbildung von 18-Krone-6 mit Alkalimetall-Kationen (Kapitel 2) sowie die Komplexierung molekularer Ammonium-Ionen und Guanidin-Derivate mit 15-Krone-5 und 18-Krone-6 (Kapitel 3) in polaren Lösungsmitteln angewandt.

Dazu wurden die ungebundenen Reaktionspartner zunächst in einem Abstand von 20 Å positioniert und die Reaktionskoordinate in 200 Reaktionsfenster zu je 0.1 Å unterteilt. Ein Reaktionsfenster beinhaltete dabei das statistisch relevante *Sampling* der Energieänderung zufällig erzeugter Konformere für einen Schritt entlang der Reaktionskoordinate. Der Erwartungswert der Änderung der Freien Energie für einen Reaktionsschritt, also die gegenseitige Annäherung von Wirt und Gast um 0.1 Å, konnte mit thermodynamischer Störungstheorie[41,42,56] berechnet werden. Durch Summation dieser inkrementellen Beiträge entlang der Reaktionskoordinate wurden radialsymmetrische potentials-of-mean-force[47] erhalten, welche formal die Ableitung der Freien Komplexbildungsenthalpie nach der Abstandskoordinate *r* darstellen. In allen betrachteten Fällen finden sich im Bereich des intermolekularen Kontaktabstandes deutliche Minima, die wohldefinierte Bindungspositionen anzeigen. Die untersuchten Metallkation-Kronenether-Komplexe zeichnen sich dadurch aus, daß die Minima der PMF bei den größeren Metallkationen K^+, Rb^+ und Cs^+ merklich vom Massenschwerpunkt des Kronenethers

abweichen. Da dieser Effekt in Methanol ausgeprägter ist als in Acetonitril, welcher die Kationen schwächer solvatisiert, spielen dabei neben sterischen Effekten auch Solvenseffekte eine wesentliche Rolle.

Die Anwesenheit eines Solvens bedingt weiterhin das Auftreten von Barrieren, die auf das Abstreifen der Solvatschalen bei der Komplexierungsreaktion zurückgeführt werden können. Während die Höhe der Barrieren in Acetonitril und Methanol unter 1 kcal/mol bleibt, ergeben sich bei den betrachteten Komplexbildungen in Wasser etwa 4 kcal/mol, bedingt durch die starken Wasserstoffbrücken.

Das Minimum eines PMF stellt für sich genommen noch kein Maß für die Freie Bindungsenthalpie dar, da es nur den günstigsten von vielen möglichen Bindungszuständen repräsentiert. Anhand der potentials-of-mean-force können allerdings Erwartungswerte der Freien Komplexbildungsenthalpie der einzelnen Komplexe bestimmt werden. Dazu wurde radial über die bindenden Zustände im Bereich eines Minimums integriert und eine notwendige Korrektur auf den Standardzustand 1 mol/l durchgeführt.[57] Weitergehende Korrekturen umfaßten die Transformation von Translations- und Rotationsfreiheitsgraden bei der Komplexbindung auf der Basis der Rotationssymetriezahlen[58] der ungebundenen und der gebundenen Reaktionspartner. Es konnte dabei gezeigt werden, daß die auf diese Weise berechneten, mittleren Freien Komplexbildungsenthalpien der Kronenether-Komplexe in guter Übereinstimmung mit experimentell bestimmten Daten[59,60,61] stehen. Während die absoluten Komplexbildungsenthalpien durch den RISM/MC Algorithmus tendenziell leicht überschätzt wurden, stimmen die beobachteten Trends innerhalb der Gruppen der Metallkationen, Ammonium-Verbindungen und Guanidin-Derivate gut mit dem Experiment überein. Eine Übersicht über die im Rahmen der Arbeit berechneten freien Bildungsenthalpien im Vergleich mit experimentellen Ergebnissen findet sich in Tabelle 1.

TABELLE 1: Berechnete freie Bindungsenthalpien ΔG und $\Delta G°$ (vgl. Kap. 2 und 3) und experimentelle Referenzwerte ΔG_{exp} im Vergleich.

Lösungsmittel	Komplex	ΔG / kcal mol^{-1}	$\Delta G°$ / kcal mol^{-1}	ΔG_{exp} / kcal mol^{-1}
H_2O	18-Krone-6/NH_4^+	-0,16	-3,43	-1,4 ± 0,2[60]
H_2O	18-Krone-6/$MeNH_3^+$	-0,6	-3,04	-0,8 ± 0,4[60]
H_2O	18-Krone-6/$EtNH_3^+$	0,0	-2,44	-0,2[60]
H_3COH	18-Krone-6/Na^+	-3,88	-8,26	-6,4 ± 1,0[59]
H_3COH	18-Krone-6/K^+	-4,28	-8,66	-8,3 ± 0,2[59]
H_3COH	18-Krone-6/Rb^+	-3,97	-8,35	-7,4 ± 0,4[59]
H_3COH	18-Krone-6/Cs^+	-3,87	-8,25	-6,4 ± 0,2[59]
H_3COH	18-Krone-6/$MeNH_3^+$	-5,05	-7,49	-5,45[60]
H_3COH	18-Krone-6/$EtNH_3^+$	-5,74	-8,18	-5,86[60]
H_3COH	18-Krone-6/$(H_2N)_2CO$	-4,65	-7,49	-3,34[61]
H_3COH	18-Krone-6/$(H_2N)_2CNH_2^+$	-4,90	-7,98	-5,80[61]
H_3COH	18-Krone-6/$(Me_2N)_2CNH_2^+$	-4,55	-7,39	<-6,81[61]
H_3COH	18-Krone-6/$PhC(NH_2)_2^+$	-4,87	-7,31	-5,16[61]
H_3COH	15-Krone-5/$(H_2N)_2CO$	-4,68	-6,87	-3,34[61]
H_3COH	15-Krone-5/$(H_2N)_2CNH_2^+$	-4,64	-7,08	-5,01[61]
H_3COH	15-Krone-5/$(Me_2N)_2CNH_2^+$	-4,07	-6,26	-3,46[61]
H_3COH	15-Krone-5/$PhC(NH_2)_2^+$	-4,97	-6,76	-4,39[61]
H_3CCN	18-Krone-6/Na^+	-8,59	-12,96	-5,9 ± 0,8[60]
H_3CCN	18-Krone-6/K^+	-8,13	-12,50	-8,0 ± 0,5[60]
H_3CCN	18-Krone-6/Rb^+	-6,80	-11,17	-6,8 ± 0,4[60]
H_3CCN	18-Krone-6/Cs^+	-6,27	-10,64	-6,6 ± 0,2[60]

6.5 Globale Optimierung

Im Hinblick auf die computergestützte Optimierung stereoselektiver Trennungen wurde ein zuverlässiges und leistungsfähiges Verfahren zur Modellierung chiraler Selektion auf der Basis automatisierten *Dockings* mit einem genetischen Algorithmus[50] entwickelt. Das Prinzip des Verfahrens besteht in der globalen Optimierung codierter struktureller Information zur Auffindung einer Population von Energieminima, ohne im Einzelfall Kenntnisse der Bindungsstelle oder der Bindungsgeometrie in das Verfahren einzubringen. Zur Reduktion der benötigten Rechenzeit wurde ein zweistufiges Verfahren gewählt. Am Anfang erfolgt eine Präoptimierung starrer Bindungspartner und, darauf folgend, eine Feinoptimierung der resultierenden Population.

Der genetische Algorithmus benutzt natürliche Evolutionsprozesse zur Entwicklung einer konstanten Population von Genen, die im Wesentlichen binärcodierte Strukturinformationen eines Komplexes beinhalten.[62] Die Population wird zunächst durch zufällige Variation der intermolekularen Freiheitsgrade erzeugt. Die darauf folgenden Generationswechsel gliedern sich in die folgenden Schritte

- Selektion geeigneter Elternpaare
- Anwendung der genetischen Operatoren zur Erzeugung von Tochtergenen
- Selektion der überlebenden, stärksten Individuen (Komplexstrukturen)

Die Selektionsschritte sind an ein geeignetes Auswahlkriterium, die *Fitneßfunktion*, gebunden, das die Gesamtenergie eines individuellen Komplexes angibt. Die Zuverlässigkeit dieser Fitneßfunktion ist dabei von hoher Bedeutung für die Ergebnisse der Optimierung. Eine geeignete Fitneßfunktion wurde durch Kombination des TRIPOS-Kraftfelds[63] mit der Integralgleichungsmethode auf der Basis von Gleichung 5.2 erhalten. Die freie Lösungsenthalpie der Komplexpartner in gebundenen und ungebundenen Zustand wurde mit Hilfe des RISM Modells berechnet.

Erste Untersuchungen mit dem hybriden RISM/GA Algorithmus zeigten, daß das Verfahren prinzipiell zum automatisierten Docking in Lösung geeignet ist. Problematisch erschien jedoch die bekannt langsame Konvergenz des einfachen genetischen Algorithmus.[50] Es erschien daher erfolgversprechend, den Suchraum des genetischen Algorithmus auf den Bereich der lokalen Minima einzuschränken, also

lokale Optimierung untergeordnet im Rahmen der globalen Optimierung für jede Komplexstruktur durchzuführen (RISM/LGA). Dabei zeigte sich, daß ein Mehraufwand an Rechenleistung bedingt durch lokale Optimierung im Rahmen des LGA dem schnelleren und zuverlässigeren Konvergenzverhalten des LGA gegenübersteht. In der Literatur sind vergleichbare Methoden,[64,65] die genetische Algorithmen mit lokale Suche verbinden, unter der Bezeichnung ‚Lamarckscher Algorithmus'[31] bekannt geworden.

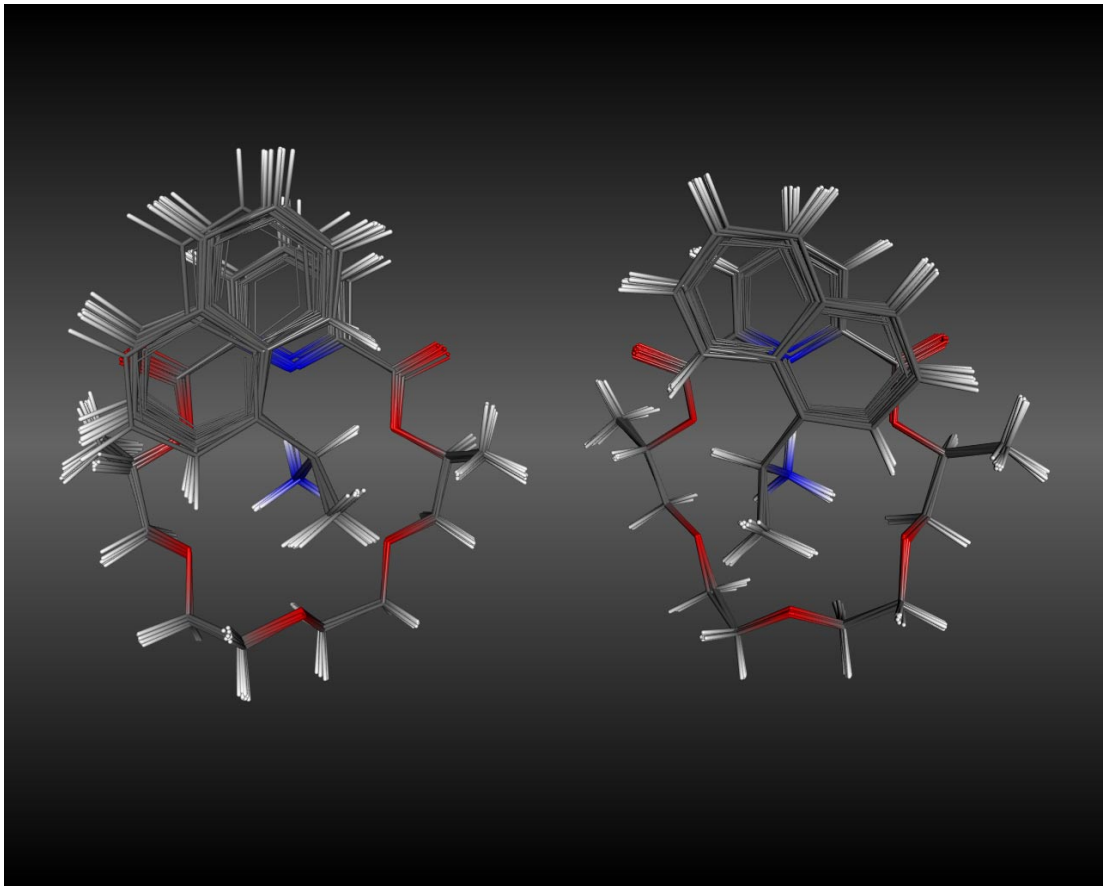

Abbildung 5: Stereoselektive Erkennung von α-(1-Naphtyl)-ethylammonium durch (S,S)-dimethyldiketopyridino-18-Krone-6. Jeweils 10 relevante Minima vom (R)-Gast (links) und (S)-Gast (rechts) sind überlagert. Die Bevorzugung des (R)-Gastes bei der Komplexierung[66] entsteht durch konformationelle Einschränkungen und sterische Wechselwirkungen der Naphtyl-Gruppe mit einem der Stereozentren des Wirtsmoleküls.

Dieses Verfahren wurde zum Docking chiraler organischer Ammoniumverbindungen mit einer Gruppe chiraler Kronenether mit Pyridin-Funktion im Lösungsmittel Methanol eingesetzt. Für jeden Komplex ergab sich nach der Präoptimierung eine Population von zehn energieminimierten Zielstrukturen. Im Rahmen der Feinoptimierung erfolgte darauf aufbauend eine vollflexible Relaxation der intramolekularen Freiheitsgrade. In allen untersuchten Fällen wurden hierbei plausible Zielstrukturen erhalten, die sich durch drei intermolekulare Wasserstoffbrückenbindungen auszeichnen. Durch die Fixierung der Wasserstoffbrücken ist der Torsionsraum der an der Bindung nicht beteiligten Seitenketten mehr oder weniger stark eingeschränkt.

Chirale Selektion entsteht durch unterschiedliche sterische Wechselwirkung der Seitenketten der Komplexpartner bei den isomeren Komplexen (vgl. Abb. 5). In der Regel sind die bei der Enantiodifferenzierung experimentell bestimmten Zahlenwerte der Energie sehr klein im Vergleich mit freien Komplexbildungsenthalpien. Zu deren Quantifizierung wurde zunächst eine Zustandssumme der zehn lokalen Minima jedes Komplexes berechnet. Anhand einer Boltzmann-Statistik konnten mittlere Energien ΔU der Komplexe und anhand des thermodynamischen Mittels freie Energien ΔA der Komplexe berechnet werden, die im vorliegenden System eine gute Approximation für Freie Bindungsenthalpien darstellen. Die resultierenden freien Bildungsenthalpien isomerer Komplexe wurden voneinander subtrahiert, und auf diesem Weg die chirale Selektion $\Delta\Delta G_{\text{R-S}}$ als Differenz der stereoisomeren Komplexe berechnet. Eine Übersicht über die im Rahmen der Studie berechneten thermodynamischen Parameter im Vergleich mit experimentellen Daten findet sich in Tabelle 2.

Aus den Daten ergibt sich, daß die berechnete Tendenz der Selektion - also die Bevorzugung von S-Isomer oder R-Isomer - im Rahmen der Studie immer richtig wiedergegeben wurde. Die Absolutwerte der berechneten Selektion $\Delta\Delta G_{\text{R-S}}$ zeigen Abweichungen von den experimentellen Ergebnissen $\Delta\Delta G_{\text{exp}}$, die teilweise innerhalb der Meßgenauigkeit liegen. Insgesamt aber wird die Selektion von der eingesetzten Methodik etwas überschätzt.

Es fällt auf, daß die Differenz der im Rahmen des RISM/LGA Verfahrens berechneten globalen Minima $\Delta W_{\text{R-S}}$ bereits recht gut die experimentell bestimmte, chirale Selektion[16] wiedergibt. Dieser Befund wird auf die allgemein geringe konformationelle Freiheit der Gastmoleküle zurückgeführt, die jeweils durch drei äqui-

valente, starke Wasserstoffbrückenbindungen fest mit ihren Wirtsmolekülen verbunden sind.

TABELLE 2: Thermodynamische Parameter diastereomerer Wirt-Gast Komplexe in kcal mol^{-1}. Die Notation der untersuchten Wirt- und Gastmoleküle ist in Kapitel 4 angegeben.

Wirt-Gast	$\Delta W_{R\text{-}S}^{rigid}$ / kcal mol^{-1}	$\Delta W_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta U_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta(TS_{R\text{-}S})$ / kcal mol^{-1}	$\Delta\Delta G_{R\text{-}S}$ / kcal mol^{-1}	$\Delta\Delta G_{exp}$ [16] / kcal mol^{-1}
2-G1	0,385	1,384	1,602	0,093	1,695	0,33
7-G2	0,748	1,307	1,205	0,407	1,611	0,19
5-G1	0,718	1,978	1,831	-0,068	1,763	0,17
4-G1	0,001	1,786	1,649	-0,289	1,360	0,58
6-G1	0,622	0,653	0,547	-0,353	0,194	0,36
3-G1	0,777	1,923	1,857	0,001	1,858	0,58
3-G2	-0,704	2,357	2,378	0,128	2,506	0,30
3-G3	0,351	0,283	0,285	0,166	0,451	0,33
3-G4	0,916	0,102	0,369	0,003	0,372	-
3-G5	0,385	0,619	0,738	0,031	0,769	-

6.6 Ausblick

Im Rahmen der vorliegenden Arbeit konnte gezeigt werden, daß problemangepaßte Hybridverfahren aus Simulationsmethoden und der Integralgleichungsmethode RISM zur computergestützten Untersuchung der Komplexbildungsthermodynamik gut geeignet sind. RISM ermöglicht die Darstellung bestimmter lokaler Eigenschaften von Energie-Hyperflächen, wie beispielsweise Minimum-Energie-Pfade, die mit anderen Methoden nur schwer zugänglich sind. Im Vergleich zur klassischen Simulation expliziter Lösungsmittelteilchen reduziert die Verwendung eines statistischen Solvensmodells den Rechenaufwand erheblich, wobei allerdings jegliche Information über die Dynamik des Lösungsmittels verloren geht. Im Rah-

men der untersuchten Komplexe konnte die benötigte Rechenzeit um bis zu einer Größenordnung verringert werden. Trotzdem ermöglichen die Verfahren die Bestimmung freier Komplexbildungsenthalpien und Selektivitäten in guter Übereinstimmung mit dem Experiment. Dies gilt insbesondere auch für differentielle Größen wie zum Beispiel chirale Selektionen.

Die Erweiterung der bestehenden Verfahren zur Anwendung auf größere Systeme erscheint sehr vielversprechend, allerdings steigt der Speicherplatzbedarf aufgrund des Skalierungsverhaltens der RISM-Matrixgleichungen quadratisch mit der Anzahl der Wechselwirkungszentren an. Bei der Untersuchung großer Biomoleküle ist daher mit technischen Schwierigkeiten zu rechnen.

Auch der Einsatz der Integralgleichungsmethode als effektives Lösungsmittelpotential im Rahmen molekulardynamischer Simulation birgt viel Entwicklungsmöglichkeiten. Letztendlich können Untersuchungen mit der Integralgleichungsmethode und der molekulardynamischen Simulation viel zum Verständnis methodischer Fehler beider Verfahren und deren Kompensation beitragen.

6.7 Literatur

(1) Lehn, J.-M. Supramolecular Chemistry: Concepts and Perspectives, VCH, Weinheim **1995**.

(2) Westman, G. Selectivity in Host-Guest Chemistry, Thesis, Department of Organic Chemistry, Chalmers **1995**.

(3) Vögtle, F. Supramolekulare Chemie, Teubner Verlag, Wiesbaden **1992**.

(4) Lide, D.R.; Ed.; CRC Handbook of chemistry and physics, CRC Press **1994**.

(5) Eliel, E.L.; Wilen, S.H.; Mander, L.N. Stereochemistry of organic compounds, Wiley Interscience New York, **1994**.

(6) Rudolph, J. Selektive Erkennung von Molekülen und stereotypen funktionellen Gruppen, Dissertation, Bochum **1995**.

(7) Ojima, I. Catalytic asymetric synthesis, VCH New York, **1993**.

(8) Arya, P.; Panda, G.; Rao, N.V.; Alper, H.; Bourque, S.C.; Manzer, L.E. *J. Am. Chem. Soc.* **2001**, *123*, 2889.

(9) Ferguson, C.G.; Thatcher, G.R.J. *Org. Letters* **1999**, *1*, 829-832.

(10) Fischer, E. *Chem. Ber.* **1894**, *27*, 2993.

(11) Horváth, G.; Huszthy, P.; Szarvas, S.; Szókán, G.; Redd, J.T.; Bradshaw, J.S.; Izatt, R.M. *Indust. Eng. Chem. Research* **2000**, *39*, 3576.

(12) Deng,Y.; Zhang, J.; Tsuda, T.; Yu, P.H.; Boulton, A.A.; Cassidy, R.M. *Anal. Chem.* **1998**, *70*, 4586.

(13) Roussel, C. Chiralchrom: A list for Chiral Separations and Chiral Discrimination, Archives available at: http://listes.cru.fr/arc/chiralchrom@cru.fr/

(14) Pirkle, W.H.; Pochapsky, T.C. *Chem. Rev*. **1989**, *89*, 347.

(15) Stoddard, J.F.; Eliel, E.L.; Wilen, S.H. Topics in stereochemistry, Wiley Interscience New York, **1988**.

(16) Zhang, X.X.; Bradshaw, J.S.; Izatt, R.M. *Chem. Rev.* **1997**, *97*, 3313.

(17) Takahashi, K.; Hattori, K. *J. Inclus. Phenom. Mol. Recog. Chem.* **1994**, *17*, 1.

(18) Lipkowitz, K.B. *Chem. Rev.* **1998**, *98*, 1741.

(19) Exner, T. Computergestützte Strukturbestimmung biochemischer Komplexe durch einen Fuzzy Logic-basierten Algorithmus, Dissertation, Darmstadt **2000**.

(20) Lee, B.; Richards, F.M. *J. Mol. Biol.* **1971**, *55*, 379.

(21) Richards, F.M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151.

(22) Dove, S.; Buschauer, A. *Quant. Struct. Act. Relat.* **1997**, *16*, 11.

(23) Schaper, K.J. *Prog. Clin. Biol. Res.* **1989**, *291*, 41.

(24) Kubinyi, H.; Ed.; 3D-QSAR in Drug Design, Escom, Leiden, **1993**.

(25) Cramer, R.D. III; Patterson, D.E.; Bunce, J.D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.

(26) Luke, B.T. *J. Chem. Inf. Comput. Sci.* **1994**, 1279.

(27) Angeline, P.J.; Reynolds, R.G.; McDonnell, J.R.; Eberhart R.C.; Eds.; Evolutionary Programming VI, Proceedings. Lecture Notes in Computer Science, Vol. 1213, Springer, **1997**.

(28) Schultz, A.C. The Genetic Algorithms Archive, The Navy Center for Applied Research in Artificial Intelligence, http://www.aic.nrl.navy.mil/galist/

(29) Jones, G.; Willett, P.; Glen, R.C. *J. Comp.-Aided Mol. Design* **1995**, *9*, 532.

(30) Westhead, D.R.; Clark, D.E.; Murray, C.W. *J. Comp.-Aided Mol. Design* **1997**, *11*, 209.

(31) Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. *J. Comp. Chem.* **1998**, *19*, 1639.

(32) Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759.

(33) Jones, D. *J. Mol. Biol.* **1997**, *267*, 727.

(34) Rarey, M.; Kramer, B., Lengauer, T., Klebe, G. *J. Mol. Biol.* **1996,** *261*, 470.

(35) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.

(36) Kirkwood, J. *J. Chem. Phys.* **1935**, *3*, 300.

(37) van Gunsteren, W.F.; Weiner, P.K. Computer simulation of biomolecular systems, Escom Leiden, **1989**.

(38) Jorgensen, W.L.; Blake, J.F.; Buckner, J.K. *J. Chem. Phys*. **1989**, *129*, 193.

(39) Zwanzig, R.W. *J. Chem. Phys.* **1954**, *22*, 1420.

(40) Denti, T.Z.M.; van Gunsteren, W.F.; Diederich, F. *J. Am. Chem. Soc*. **1996**, *118*, 6044.

(41) Pranata, J.; Jorgensen, W.L. *Tetrahedron* **1991**, *47*, 2491.

(42) Nguyen, T.B.; Jorgensen, W.L. *Proc. Natl. Acad. Sci.* **1993**, *90*, 1194.

(43) Marszalek, P.; Lu, H.; Li, H.; Carrion, M.; Oberhauser, A.; Schulten, K.; Fernandez, J. *Nature* **1999**, *402*, 100.

(44) Chandler, D.; Andersen, H.C. *J. Chem. Phys.* **1972**, *57*, 1930.

(45) Hirata, F.; Rossky, P.J. *Chem. Phys. Lett.* **1981**, *83*, 329.

(46) Lee, P.H.; Maggiora, G.M. *J. Phys. Chem.* **1993**, *97*, 10175.

(47) Allen,M.P.; Tildesley, D.J. Computer Simulation of Liquids, Claredon Press, Oxford **1987**.

(48) Singer, S.J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621.

(49) Numerical Recipes in C: THE ART OF SCIENTIFIC COMPUTING, Cambridge University Press, **1988**, 408.

(50) Carroll, D.L. PROGRAM GAFORTRAN (version 1.7), University of Illinois, 306 Talbot Lab, 104 S. Wright St., Urbana, IL 61801, **1998**.

(51) Karplus, M. CHARMM Documentation, The CHARMM Development Project, Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

(52) Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C. D. In *Data Visualization in Molecular Science - Tools for Insight and Innovation*; Bowie, J. E.; Ed.; Addison-Wesley: Reading **1995**.

(53) Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G. *Molecular Models: Visualization.* In: Schleyer, P.v.R.; Allinger, N.L.; Clark, T.;

Gasteiger, J.; Kollman, P.A.; Schaefer III, H.F.; Schreiner, P.R.; Eds.; "The Encyclopedia of Computational Chemistry", John Wiley & Sons: Chichester, **1998**.

(54) Kowall, T.; Geiger, A. *J. Phys. Chem.* **1995**, *99*, 5240.

(55) Metropolis N. Journal Chem. Phys. **1953**, *21*, 1087.

(56) Prue, J.E. *J. Chem. Educ.* **1969**, *46*, 12.

(57) Gilson, M.K.; Given, J.A.; Bush, B.L.; McCammon, J.A. *Biophys. J.* **1997**, *72*, 1047.

(58) Hill, T.L. Statistical Thermodynamics, Addison-Wesley Publishing, London **1950**.

(59) Izatt, R. M.; Bradshaw, J. S.; Nielsen, S. A.; Lamb, J. D.; Christensen, J. J. Chem. Rev. **1985**, *85*, 271.

(60) Izatt, R.M.; Pawlak, K.; Bradshaw, J.S.; Bruening, R.L. Chem. Rev. **1995**, *95*, 2529.

(61) Buschmann, H.-J-; Dong, H.; Mutihac, L.; Schollmeyer, E. J. Therm. Anal. Cal. **1999**, *57*, 487.

(62) Heistermann, J. Genetische Algorithmen - Theorie und Praxis evolutionärer Optimierung, B.G. Teubner Verlagsgesellschaft, Stuttgart **1994**.

(63) SYBYL 6.6, *Tripos Inc.*, 1699 South Hanley Road, St. Louis, Missouri, 63144 USA.

(64) Gottlieb, J. Evolutionary Algorithms for Constrained Optimization Problems, Dissertation, Institut für Informatik, Technische Universität Clausthal **1999**.

(65) Freisleben, B.; Merz, P. Proc. IEEE International Conference on Evolutionary Computation, IEEE, World Congress on Computational Intelligence, **1996**, 616.

(66) Dearden, D.V.; Dejsupa, C.; Liang, Y.; Bradshaw, J.S.; Izatt, R.M. J. Am. Chem. Soc. **1997**, *119*, 353.

Hilfsmittel

Folgende Hilfsmittel wurden im Rahmen dieser Arbeit eingesetzt:

Die dreidimensionalen Molekülstrukturen (Eingabedaten) wurden mit dem Computerprogramm *Sybyl* (Version 6.7, Tripos Inc., 1699 South Hanely Road, St. Louis, Missouri, 63144, USA) modelliert.

- Zur Parametrisierung der bindenden und nichtbindenden Wechselwirkungen wurden die OPLS Kraftfeldparameter (W.L. Jorgensen, BOSS Version 4.1, Yale University, New Haven, Connecticut, 2000) und TRIPOS Kraftfeldparameter (Tripos Inc., 1699 South Hanely Road, St. Louis, Missouri, 63144, USA) verwendet.

- Für die thermodynamischen Berechnungen diente eine SGI Power Challenge, 10 MIPS-Prozessoren R10000, 190 MHz und drei Intel Pentium III, 0.5-1 GHz.

- Ein genetischer Algorithmus wurde auf der Basis des Programms GaFortran 1.7 (D.L. Carroll, University of Illinois, 306 Talbot Lab, 104 S. Wright St., Urbana, IL 61801, USA) entwickelt.

- Die schriftliche Ausarbeitung wurde mit dem Textverarbeitungsprogramm Word 97 (Microsoft Inc.) vorgenommen. Darin enthaltene Tabellen wurden mit dem Tabellenkalkulationsprogramm Excel 97 (Microsoft Inc.) erstellt.

- Graphen wurden mit dem Programm *Origin* 3.5 (Microcal Software Inc.) und Matlab (Student Edition 12, The MathWorks Inc.) dargestellt.

- Molekülvisualisierungen erfolgten mit Hilfe des Programms *Molcad II* (Version 1.3, M. Keil, J. Brickmann, Physikalische Chemie I, Petersenstr. 20, 64287 Darmstadt).

- Weitere Abbildungen wurden mit Isis Draw 2.1.3d (MDL Information Systems Inc.) und Designer 7.1 (Micrografx Inc) erstellt.

Friedemann Schmidt
Arheilger Str. 94
64289 Darmstadt

Darmstadt, den 19. Dezember 2001

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, daß ich meine Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

_____

(Friedemann Schmidt)

Lebenslauf

Friedemann Schmidt

Geburtsdatum:		07.01.1972
Geburtsort:		München
Familienstand:		ledig
Schulbildung:	09/78 - 08/82	Grundschule Riemerling 85521 Riemerling
	09/82 - 07/91	Gymnasium Ottobrunn 85521 Ottobrunn
Studium:	04/92 - 09/97	Chemiestudium an der Technischen Universität Darmstadt
	09/95 - 03/96	Chemiestudium an der Universität Zürich „Quantum beat spectroscopy of NO_3“
	02/97 - 06/97	Diplomarbeit bei Prof. Dr. J. Brickmann: „Analytische Berechnung elektrostatischer Solvatationsenergien von Molekülen“
Promotion:	10/97 - 09/01	Doktorand bei Prof. Dr. J. Brickmann: „Fast computational treatment of molecular complexation in solution: Hybrid integral equation approaches“